A DYNAMIC RESPONSE SURFACE MODEL
FOR FREQUENCY DOMAIN
SIMULATION EXPERIMENTS

by

Lee W. Schruben
David Heath
Arnold Buss*

* The John M. Olin School of Business, Washington University, St. Louis, MO. 63130
This work was sponsored in part by National Science Foundation Grant ECS 8810517 and by subcontract B-F0448-A-E from Battelle Pacific Northwest Laboratories. This support is gratefully acknowledged.
# TABLE OF CONTENTS

0. INTRODUCTION................................................................. 0.1

1. A DYNAMIC RESPONSE SURFACE MODEL........................................ 1.1

2. FREQUENCY DOMAIN EXPERIMENTS FOR FACTOR SCREENING.................. 2.1
   2.1. TERM INDICATOR FREQUENCIES........................................... 2.1
   2.2. SELECTION OF DRIVING FREQUENCIES..................................... 2.14
   2.3. PERIODIC STEADY STATE.................................................. 2.24
   2.4. SYNCHRONOUS VS ASYNCHRONOUS SIMULATIONS........................... 2.25

3. EXAMPLE OF FREQUENCY DOMAIN EXPERIMENTS FOR FACTOR SCREENING....... 3.1

4. ESTIMATING THE DC RESPONSE............................................... 4.1
   4.1. THE KRAMERS-KRONIG RELATIONS......................................... 4.1
   4.2. EXAMPLE: M/M/1 QUEUE.................................................. 4.8
   4.3. DISCUSSION.............................................................. 4.9

5. PARAMETRIC ESTIMATION OF THE IMPULSE RESPONSE FUNCTION.................. 5.1
   5.1. THE ESTIMATOR.......................................................... 5.1
   5.2. AN EXAMPLE: M/M/∞ QUEUE.............................................. 5.4
   5.3. FREQUENCY SELECTION.................................................... 5.7

REFERENCES........................................................................... R.1

APPENDIX I. GENERALIZED FUNCTIONS AND PROCESSES............................. I.1
   I.1. HEURISTIC TREATMENT OF THE DIRAC DELTA FUNCTION.................. I.1
   I.2. TEST FUNCTIONS........................................................... I.4
   I.3. TEMPERED DISTRIBUTIONS................................................. I.5
   I.4. FOURIER TRANSFORM OF TEMPERED DISTRIBUTIONS..................... I.7
   I.5. DIFFERENTIATION OF TEMPERED DISTRIBUTIONS......................... I.11
   I.6. GENERALIZED PROCESSES.................................................. I.12
LIST OF FIGURES

FIGURE 1.1 A CONVENTIONAL MODEL OF SIMULATION.................................1.4
FIGURE 1.2 PARAMETERS AS INPUTS......................................................1.4
FIGURE 2.1 DENOMINATOR=27 BY FOLDING AT 14½..................................2.22
FIGURE 2.2 DENOMINATOR=28 BY FOLDING AT 16..................................2.22
FIGURE 2.3 STARTING HEURISTIC WITH NUMERATOR = 2..........................2.22
FIGURE 2.4 STARTING HEURISTIC WITH NUMERATOR = 4..........................2.22
FIGURE 2.5 DRIVING NUMERATORS = \{1,5,8\}......................................2.22
FIGURE 3.1 FREQUENCY RESPONSE "NOISE RUN"..................................3.6
FIGURE 3.2 FREQUENCY RESPONSE.....................................................3.7
FIGURE 3.2 FREQUENCY RESPONSE.....................................................3.8
FIGURE 3.4 FREQUENCY RESPONSE.....................................................3.9
LIST OF TABLES

TABLE 2.1 TERM INDICATOR FREQUENCIES FOR THE QUADRATIC MODEL..............2.13
TABLE 2.2 HEURISTICS FOR FREQUENCY SELECTION.................................2.20
TABLE 2.3 FREQUENCIES FOR VARIOUS NUMBERS OF FACTORS.....................2.21
TABLE 3.1 TERM INDICATOR FREQUENCIES: THREE PARAMETERS, SECOND ORDER....3.5
TABLE 4.1 KRAMERS-KRONIG PROCEDURE APPLIED TO M/M/1 QUEUE...............4.10
0. Introduction

Frequency domain experiments were introduced by Schruben and Cogliano (1987) as an efficient method for screening factors in building a response surface model of a simulation. In this paper we extend and formalize their approach by introducing a meta-model, the dynamic response surface model. This model attempts to capture two properties of simulation models which make their analysis difficult: non-linearities and memory. We assume that the non-linearities may be approximated by a polynomial and the memory by a subsequent linear filter. We further assume that randomness in the system is additive but not necessarily uncorrelated. Narendra and Gallman (1966) present a deterministic version of this model in discrete time and refer to it as the Hammerstein model.

Frequency domain experiments can obtain useful information about the sensitivity of the simulation to a number of factors. The benefit of this approach is that relatively few runs are needed. A conventional experimental design may require \( O(k^p) \) runs, where \( p \) is the number of factors and \( k \) is one greater than the order of the polynomial response surface. Frequency domain screening experiments typically require \( << O(p) \) runs, albeit with somewhat longer run lengths.

In a frequency domain experiment, factors which are ordinarily held fixed are varied in a systematic (or periodic) fashion throughout the run. The frequency content of the output is examined to determine influential terms. Peaks in the output power spectrum correspond to influential terms, while the absence of peaks indicates non-influential terms. If the frequencies are chosen carefully, potential peaks for distinct terms will be at distinct frequencies.
The organization of this paper is as follows. Section 1 introduces the
dynamic response surface model. Section 2 presents the frequency domain
approach to simulation experiments and discusses design and implementation
issues, including the problem of selecting the appropriate frequencies. In
section 3 some examples illustrate the use of frequency domain experiments.
Section 4 presents a different approach based on the relationship between the
real and imaginary parts of certain Fourier transforms, the Kramers-Kronig
relations. Section 5 deals with fitting a model to a simulation by assuming
a particular form for the impulse response functions. An appendix deals with
some formalities concerning generalized functions and processes.
1. A Dynamic Response Surface Model

Simulations of systems are often performed when analytical methods yield no closed-form solutions. The simulation is often highly complex. A model of the simulation, or meta-model, may be used to draw inferences on the behavior of the simulation model as a "black box" (Kleijnen, 1979). In this section we present one such model, the Dynamic Response Surface Model (DRSM). This model facilitates frequency domain analysis, as discussed in the following section.

A common view of simulation is as follows: A stream of inputs, usually generated pseudo-randomly, are transformed by the simulation into a stream of outputs. Parameters of the model may be associated with the inputs (e.g. mean arrival rates to a queue) or with the system model (e.g. number of machines in a factory). Randomness is associated with both the inputs and with the system. This meta-model of a simulation can be represented as in Figure 1.1.

When they introduced frequency domain simulation experiments, Schruben and Cogliano (1981, 1987) suggested a departure from conventional approaches to simulation, that of viewing parameters as inputs to the system. In run-oriented simulation experiments, parameters are held fixed during the the run. However, for frequency domain simulation experiments they are oscillated throughout the run. Any reasonable meta-model must reflect the time-varying of parameters. This is illustrated in Figure 1.2.

The Dynamic Response Surface Model (DRSM) makes explicit assumptions concerning the structure of ψ which incorporate both the nonlinearity of ψ in x(t) and the memory of ψ. The DRSM models the non-linearity as a polynomial in x(t). The values of each polynomial term are then passed through a linear filter. The terms in the polynomial are called "pseudo-linear" since, as will
be seen, they correspond to sets of factors each of which behaves (in the frequency domain) as if it were a linear term oscillated at different frequencies, the term indicator frequencies (see Section 2).

It is convenient to introduce multi-index notation. Let \( \mathbf{x} \) be a vector in \( \mathbb{R}^n \) and \( \alpha \) a vector in \( \mathbb{Z}_+^n \) (\( \mathbb{Z}_+ \) = the set of non-negative integers). The term \( \mathbf{x}^\alpha \) is defined to be

\[
\frac{1}{n} \sum_{j=1}^{n} [x_j]^{\alpha_j}
\]

where \( x_j \) and \( \alpha_j \) are the components of \( \mathbf{x} \) and \( \alpha \) respectively. The vector \( \alpha \) is called a multi-index. The modulus \( |\alpha| \) of a multi-index \( \alpha \) is defined to be

\[
|\alpha| = \alpha_1 + \ldots + \alpha_n
\]

We may now write the DRSW as follows:

\[
Y(t) = \sum_{|\alpha| \leq k} \int_0^t x^\alpha(t - \tau) g_\alpha(\tau) \, d\tau + \varepsilon(t)
\]

where \( x(t) \) are the input parameters, \( g_\alpha(\tau) \) are impulse response functions (one for each term \( x^\alpha \)), and \( \{\varepsilon(t)\} \) is a zero mean random process. The functions \( g_\alpha(\tau) \) characterize the memory of the system; there is one such function for each term in the polynomial. In some cases all \( g_\alpha \)'s will be identical, resulting in a somewhat simpler model. The process \( \{\varepsilon(t)\} \) may be viewed as "noise", and the assumption that it enters the model in an additive way is crucial to the analysis that follows. However, we do not assume it to be "white" (independent with a constant spectrum), although for practical purposes it may be assumed to be transformed (filtered) white noise. In most computer implementations of simulation models, (pseudo) random processes are generated from (pseudo) random streams of iid uniform random variables. These original streams may be thought of as white noise, and the resulting
processes are therefore a transformation of this white noise.

The other important property of this model is that the order of the polynomial, k, is specified a priori. Frequency domain experiments can be misleading or incorrect if the order of the polynomial is not sufficiently high. However, in our experience it is usually sufficient to consider second order (quadratic) models. The "response surface" part of the DRSM comes from this polynomial in \( x(t) \), and the "dynamics" are supplied by the impulse response functions \( g_x \). In the following section the DRSM will be analyzed using a frequency domain experiment.
Figure 1.1. A Conventional Meta-Model of Simulation

\( \xi_{x_I}(t) \) = input stream with (vector) parameter \( x_I \).

\( \psi_{x_S} \) = system model with (vector) parameter \( x_S \).

\( Y(t) \) = output.

Figure 2. Parameters as Inputs

\( x(t) \) = vector process of (time varying) parameters

\( \psi \) = system model

\( Y(t) \) = scalar output
2. Frequency Domain Experiments for Factor Screening

2.1. Term Indicator Frequencies

A frequency domain experiment explores a region of the parameter space of a static response surface model by analyzing the response to oscillated parameters. This analysis is based on the DRSW meta-model defined in section 1. A frequency domain experiment uses the following time-varying inputs to the DRSW (equation 1.1):

\[ x_j(t) = a_j + b_j \cos 2\pi \omega_j t \quad (j=1,\ldots,n). \]  

(2.1)

The distinct frequencies \( \omega_1, \ldots, \omega_n \) are called driving frequencies, and are chosen in such a manner that peaks in the output power spectrum may be associated with each of the different input terms. These frequencies are called term indicator frequencies (see Schruben and Cogliano, 1987) because of this association. Each term in the polynomial is of the form \( x^\alpha \) and is called a pseudo-linear term. This is because products in the time domain lead to sums and differences of corresponding frequencies in the frequency domain.

The frequency content of the output \( y(t) \) depends on the driving frequencies of the inputs, which terms are actually present in the model, and the frequency content of the noise process. Each pseudo-linear term \( x^\alpha \) can produce peaks in the output power spectrum, \( H_y(\omega) \), at frequencies

\[ \left\{ \sum_{\alpha_j > 0} (-1)^j \beta_j \omega_j : \beta \leq \alpha, |\gamma| \leq |\alpha|, \gamma_j \in \{0,1\} \right\} \]  

(2.2)

where, \( \beta \) and \( \gamma \) are multi-indices. The \( \omega_j \)'s in expression (2.2) are the term indicator frequencies for the "full" model; that is, the model in which every
possible term is present. The observation of a peak in $H_y(\omega)$ at a term
indicator frequency is evidence that the corresponding term is influential in
the output.

Equivalently, the set of term indicator frequencies is the sum of the
sets

$$\{ \pm \beta_j \omega_j : \beta_j \leq \alpha_j \}, \quad j=1,\ldots,n,$$

(2.3)

where the sum of a collection of sets is defined to be the set of all sums
having exactly one element from each set. If $a_j = 0$ for $j = 1,\ldots,n$, then the
term indicator frequencies are the sum of the sets (Schruben and Cogliano,
1987):

$$\{ (\alpha_j - 2k) \omega_j : k = 0,\ldots,\alpha_j \} \quad (j = 1,\ldots,p)$$

(2.4)

For example, suppose that $p = 2$ and $k = 2$. Assuming that $a_1 = a_2 = 0$,
the set of term indicator frequencies for $(x_1,x_2)^{(1,2)}$ consists of the sum of
the sets $\{\omega_1,\omega_2\}$ and $\{2\omega_2,-2\omega_2\}$; that is, the set

$\{\omega_1,-\omega_1,\omega_1+2\omega_2,\omega_1-2\omega_2,-\omega_1+2\omega_2,-\omega_1-2\omega_2\}.$

We now consider, for the time being, the response of a deterministic
system to a single linear term:

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

(2.5)

with the input $x(t)$ is given by

$$x(t) = a + b \cos 2\pi \omega_0 t.$$  

(2.6)
We assume that $\int g(\tau) \, d\tau < \infty$. The response is

$$y(t) = \int_0^\infty a \, g(\tau) \, d\tau + \int_0^\infty b \cos 2\pi\omega_0(t - \tau) \, g(\tau) \, d\tau$$

$$= a \int_0^\infty g(\tau) \, d\tau + \text{Re} b \int_0^\infty e^{2\pi i\omega_0(t-\tau)} g(\tau) \, d\tau$$

$$= a \hat{g}(0) + b \text{Re} e^{2\pi i\omega_0 t} \hat{g}(\omega_0)$$  \hspace{1cm} (2.7)

in which

$$\hat{g}(\omega) = \int_0^\infty e^{-2\pi i\tau} g(\tau) \, d\tau$$  \hspace{1cm} (2.8)

is the Fourier transform of $g(t)$ and "Re" denotes the real part of a complex quantity. In Equations 2.7-2.8 we have used the fact that, for $i = \sqrt{-1}$,

$$e^{i\theta} = \cos \theta + i \sin \theta,$$

so $\text{Re}(e^{i\theta}) = \cos \theta$. It will be convenient at times to write $\hat{g}(\omega)$ in polar form:

$$\hat{g}(\omega) = |\hat{g}(\omega)| \, e^{2\pi i \phi(\omega)}$$  \hspace{1cm} (2.9)

in which $|\hat{g}(\omega)|$ is the (complex) modulus of $\hat{g}(\omega)$ and $\phi(\omega)$ is its argument, or phase. If $|\phi(\omega)| < \pi$ then $\phi(\omega) = \arctan(\text{Im} \, \hat{g}(\omega) / \text{Re} \, \hat{g}(\omega))$. Although the system is deterministic, we may consider the mean and (auto) covariance using time averages (see Ljung, 1987, pp. 27-28). The time-averaged mean is
\[
m_y = \lim_{T \to \infty} \frac{1}{T} \int_0^T y(t) \, dt
= \lim_{T \to \infty} \frac{1}{T} \int_0^T dt \int_0^\infty [a + b \cos 2\pi \omega_0 (t - \tau)] \, g(\tau) \, d\tau
= \lim_{T \to \infty} \frac{1}{T} \int_0^T dt \left\{ a \int_0^\infty g(\tau) \, d\tau + b \cos 2\pi \omega_0 t \int_0^\infty \cos 2\pi \omega_0 \tau \, g(\tau) \, d\tau \right. \\
\left. + \sin 2\pi \omega_0 t \int_0^\infty \sin 2\pi \omega_0 \tau \, g(\tau) \, d\tau \right\}
= a \hat{g}(0).
\]  

Similarly, the time-averaged covariance is:

\[
R_y(t) = \lim_{T \to \infty} \frac{1}{T} \int_0^T y(s) y(s + t) \, ds
= \lim_{T \to \infty} \frac{1}{T} \int_0^T ds \int_0^\infty x(s - \tau) \, g(\tau) \, d\tau \int_0^\infty x(s + t - \eta) \, g(\eta) \, d\eta
= b^2 \int_0^\infty g(\tau) \, d\tau \int_0^\infty g(\eta) \, d\eta \\
\cdot \lim_{T \to \infty} \frac{1}{T} \int_0^T \cos 2\pi \omega_0 (s - \tau) \cos 2\pi \omega_0 (s + t - \eta) \, ds
= \frac{1}{2} b^2 \int_0^\infty g(\tau) \, d\tau \int_0^\infty g(\eta) \, d\eta \cos 2\pi \omega_0 (t + \tau - \eta)
= \frac{1}{2} b^2 G(\omega_0) \cos 2\pi \omega_0 t,
\]

where

\[
G(\omega) = |\hat{g}(\omega)|^2
\]
is the gain of the system. In obtaining the right hand side of equation 2.11 we have used Fubini's theorem and the dominated convergence theorem, together with the assumption that \( \int g(\tau) d\tau < \infty \). The power spectrum of \( \{y(t)\} \) is therefore

\[
f_y(\omega) = \int_{-\infty}^{\infty} e^{2\pi i \omega t} R_y(t) \, dt
\]

\[
= \frac{1}{4} \ b^2 \ G(\omega_0) \ \left[ \delta(\omega - \omega_0) + \delta(\omega + \omega_0) \right]
\]

(2.13)
in which \( \delta \) is the Dirac delta function (see Appendix I). The output power spectrum consists of "spikes" at frequencies \( \pm \omega_0 \), providing \( G(\omega_0) \) is not zero (see Morrice and Schruben, 1988). The presence of a peak in the estimated power spectrum at \( \pm \omega_0 \) is therefore evidence for the presence of this linear term in the model.

Now consider a term of the form \( x^p(t) \) where \( p \) is a positive integer. We have
\[
x^p(t) = (a + b \cos 2\pi\omega_0 t)^p \\
= \sum_{k=0}^{p/2} \left[ \begin{array}{c} p \\ 2k \end{array} \right] a^{p-2k} b^{2k} \cos^{2k} 2\pi\omega_0 t \\
= \sum_{k=0}^{(p-1)/2} \left[ \begin{array}{c} p \\ 2k+1 \end{array} \right] a^{p-2k-1} b^{2k+1} \cos^{2k+1} 2\pi\omega_0 t \\
\left[ \begin{array}{c} p \\ 2k+1 \end{array} \right] \sum_{k=0}^{2k+1} \left[ \begin{array}{c} 2k+1 \\ k-q \end{array} \right] \cos 2\pi(2q+1)\omega_0 t \\
+ \sum_{k=1}^{[p/2]} \left[ \begin{array}{c} p \\ 2k \end{array} \right] \left[ \begin{array}{c} 2k \\ k \end{array} \right] a^{p-2k} (b/2)^{2k} \\
= \sum_{k=0}^{2} \sum_{k=0}^{p-1} \left[ \begin{array}{c} p \\ 2k+1 \end{array} \right] a^{p-2k-1} (b/2)^{2k+1} \cos 2\pi(2q+1)\omega_0 t \\
\left[ \begin{array}{c} p \\ 2k+1 \end{array} \right] \sum_{k=0}^{2k+1} \left[ \begin{array}{c} 2k+1 \\ k-q \end{array} \right] \cos 2\pi(2q+1)\omega_0 t \\
+ \sum_{q=1}^{[p/2]} \sum_{k=1}^{k-q} \left[ \begin{array}{c} p \\ 2k \end{array} \right] \left[ \begin{array}{c} 2k \\ k+q \end{array} \right] a^{p-2k} (b/2)^{2k} \cos 2\pi(2q)\omega_0 t \\
\left[ \begin{array}{c} p \\ 2k+1 \end{array} \right] \sum_{k=0}^{2k+1} \left[ \begin{array}{c} 2k+1 \\ k-q \end{array} \right] \cos 2\pi(2q+1)\omega_0 t \\
\sum_{r=0}^{p} C_r(a,b,p) \cos 2\pi r\omega_0 t \quad (2.14) \\
\]

where
\[ C_0(a,b,p) = \sum_{k=0}^{[p/2]} \left[ \begin{array}{c} p \\ 2k \end{array} \right] \left[ \begin{array}{c} 2k \\ k \end{array} \right] a^{p-2k} (b/2)^{2k} \]

\[ C_r(a,b,p) = \sum_{k=1}^{[p/2]} 2 \left[ \begin{array}{c} p \\ 2k \end{array} \right] \left[ \begin{array}{c} 2k \\ k+q \end{array} \right] a^{p-2k} (b/2)^{2k}, \quad r = 2q, \quad q \geq 1 \]

\[ C_r(a,b,p) = \sum_{k=q}^{[p-1]/2} 2 \left[ \begin{array}{c} p \\ 2k+1 \end{array} \right] \left[ \begin{array}{c} 2k+1 \\ k+q \end{array} \right] a^{p-2k-1} (b/2)^{2k+1}, \quad r = 2q+1, \quad q \geq 0 \]

Thus, the response will contain power at all frequencies \( r\omega_0 \) for \( r \leq p \). The output power spectrum is

\[ f_y(\omega) = \sum_{r=0}^{p} C_r^2(a,b,p) G(\omega) \left[ \delta(\omega-r\omega_0) + \delta(\omega+r\omega_0) \right]. \quad (2.15) \]

For the general term \( x^{\alpha}(t) \) where \( x(t) \) is a vector and \( \alpha \) is a multi-index and

\[ x_j(t) = a_j + b_j \cos 2\pi \omega_j t \quad (2.16) \]

we have

\[ x^{\alpha}(t) = \prod_{j=1}^{n} (a_j + b_j \cos 2\pi \omega_j t)^{\alpha_j} \]

\[ = \prod_{j=1}^{n} \sum_{\lambda_j = 0}^{\alpha_j} C_{\lambda_j}^{\alpha_j} (a_j, b_j, \lambda_j) \cos 2\pi \lambda_j \omega_j t \]

\[ = \sum_{\lambda \leq \alpha} \prod_{j=1}^{n} C_{\lambda_j}^{\alpha_j} (a_j, b_j, \lambda_j) \cos 2\pi \lambda_j \omega_j t \]

where \( \lambda = (\lambda_1, \ldots, \lambda_n) \) is a multi-index and the functions \( C \) are as defined above. Applying the formula for the product of cosines, we have
\[ x^{\alpha}(t) = \sum_{\lambda \leq \alpha} \left[ \prod_{j=1}^{n} C_{\lambda}^{j}(a_j, b_j, \alpha_j) \right] 2^{-n} \sum_{\gamma} \cos 2\pi \sum_{j=1}^{n} (-1)^{\gamma_j} \lambda_j \omega_j t \] (2.17)

where the middle sum is over all multi-indices \( \gamma \) such that \( \gamma_j \in \{0,1\} \).

Observe that the frequencies in Equation (2.18) are all possible sums and differences of \( \pm \lambda_j \omega_j \), \( j = 1, \ldots, n \) and that only those with \( \lambda_j \neq 0 \) will be present. Furthermore, since cosine is an even function, we have

\[ \cos 2\pi \sum_{j=1}^{n} (-1)^{\gamma_j} \lambda_j \omega_j t = \cos 2\pi \sum_{j=1}^{n} (-1)^{1-\gamma_j} \lambda_j \omega_j t. \] (2.18)

Thus, there are two identical terms for each pair of multi-indices \( \gamma \) and \( \lambda \).

Finally, the response to \( x^{\alpha}(t) \) has stationary covariance (in the time-averaged sense)

\[ R(t) = \frac{\chi}{\lambda \leq \alpha} C_{\lambda}^{2}(a, b, \alpha) \cdot 2^{-2n} \sum_{\gamma} G\left[ \sum_{j=1}^{n} (-1)^{\gamma_j} \lambda_j \omega_j \right] \cdot \cos 2\pi \sum_{j=1}^{n} (-1)^{\gamma_j} \lambda_j \omega_j t \] (2.19)

where

\[ C_{\lambda}(a, b, \alpha) = \prod_{j=1}^{n} C_{\lambda}^{j}(a_j, b_j, \alpha_j). \]

The power spectrum is
\[
\begin{align*}
  f(\omega) &= 2^{-(n+1)} \sum_{\lambda \leq \alpha} C^2(\alpha, b, \alpha) \sum_{j=1}^{n} (-1)^{j} \lambda_j \omega_j \\
  &\cdot \left[ \delta \left( \omega - \sum_{j=1}^{n} (-1)^{j} \lambda_j \omega_j \right) \delta \left( \omega - \sum_{j=1}^{n} (-1)^{j} \lambda_j \omega_j \right) \right].
\end{align*}
\]

This verifies expression (2.2) for term indicator frequencies associated with \( x^\alpha(t) \). To verify expression (2.4) when \( a_j = 0 \), notice that

\[
(b \cos 2\pi \omega_0 t)^p = \begin{cases} 
  \left[ \frac{b}{2} \right]^p \left[ \sum_{k=1}^{p/2} \left[ \frac{p}{k} \right] \cos 2\pi(2k)\omega_0 t + \left[ \frac{p}{p/2} \right], \\
  & \text{if } p \text{ is even} \\
  \left[ \frac{b}{2} \right]^p \left[ \sum_{k=0}^{(p-1)/2} \left[ \frac{p}{k} \right] \cos 2\pi(2k+1)\omega_0 t \right], \\
  & \text{if } p \text{ is odd}
\end{cases}
\]

Thus, the term indicator frequencies for a single term in this case are all even or odd multiples of \( \omega_0 \) less than or equal to \( p \), depending on whether \( p \) is even or odd. Applying this to the general term gives expression (2.4).

To illustrate, let us consider the important case of quadratic models: those for which the polynomial in the DRSW is of degree \( k = 2 \). For this case write the model as

\[
Y(t) = \int_{0}^{\infty} g_0(\tau) \, d\tau + \sum_{j=1}^{p} \int_{0}^{\infty} x_j(t-\tau) \, g_j(\tau) \, d\tau \\
+ \sum_{j=1}^{p} \int_{0}^{\infty} x_j^2(t-\tau) \, g_{jj}(\tau) \, d\tau \\
+ \sum_{1 \leq i < j \leq p} \int_{0}^{\infty} x_i(t-\tau) \, x_j(t-\tau) \, g_{ij}(\tau) \, d\tau + \varepsilon(t).
\]

Let \( x_j(t) = a_j + b_j \cos 2\pi \omega_j t, \quad j = 1, \ldots, p \). The response is:
\[ Y(t) = \hat{\varepsilon}_0(0) + \sum_{j=1}^{p} \left[ a_j \hat{\varepsilon}_j(0) + \frac{\nu}{2} b_j^2 \hat{\varepsilon}_{jj}(0) \right] \]
\[ + \sum_{j=1}^{p} b_j |\hat{\varepsilon}_j(\omega_j)| \cos[2\pi\omega_j t + \varphi_j(\omega_j)] \]
\[ + \sum_{i,j=1}^{p} a_i b_j |\hat{\varepsilon}_{ij}(\omega_j)| \cos[2\pi\omega_j t + \varphi_{ij}(\omega_j)] \]
\[ + \frac{\nu}{2} \sum_{j=1}^{p} b_j^2 |\hat{\varepsilon}_{jj}(2\omega_j)| \cos[2\pi(2\omega_j) t + \varphi_{jj}(2\omega_j)] \]
\[ + \frac{\nu}{2} \sum_{1 \leq i < j \leq p} b_i b_j \left[ |\hat{\varepsilon}_{ij}(\omega_i + \omega_j)| \cos[2\pi(\omega_i + \omega_j) t + \varphi_{ij}(\omega_i + \omega_j)] \right] + \varepsilon(t) \]

in which \( \varphi_j \) and \( \varphi_{ij} \) are, respectively,

\[ \varphi_j(\omega) = \text{arg}(\hat{\varepsilon}_j(\omega)) \]

\[ \varphi_{ij}(\omega) = \text{arg}(\hat{\varepsilon}_{ij}(\omega)) \]

Thus, the covariance of \( Y(t) \) is

\[ R_Y(t) = \frac{\nu}{2} \left[ \sum_{j=1}^{p} b_j^2 G_j(\omega_j) + \sum_{i,j=1}^{p} a_i b_j G_{ij}(\omega_j) \right] \cos 2\pi\omega_j t \]
\[ + \frac{1}{8} \left[ \sum_{j=1}^{p} b_j^4 G_{jj}(2\omega_j) \right] \cos 2\pi(2\omega_j) t \]
\[ + \frac{1}{8} \sum_{1 \leq i < j \leq p} b_i b_j \left[ G_{ij}(\omega_i + \omega_j) \cos 2\pi(\omega_i + \omega_j) t \right] + R_\varepsilon(t) \]
and the power spectrum is

\[
f_Y(\omega) = \frac{1}{4} \left[ \sum_{j=1}^{p} b_j^2 G_j(\omega) + \sum_{i,j=1}^{p} a_i^2 b_j^2 G_{ij}(\omega) \right] \\
\cdot \left[ \delta(\omega - \omega_j) + \delta(\omega + \omega_j) \right] \\
+ \frac{1}{16} \left[ \sum_{j=1}^{p} b_j^4 G_{jj}(\omega) \right] \left[ \delta(\omega - 2\omega_j) + \delta(\omega + 2\omega_j) \right] \\
+ \frac{1}{16} \sum_{1 \leq i < j \leq p} b_i^2 b_j^2 \left[ G_{ij}(\omega) \left[ \delta(\omega - \omega_i - \omega_j) + \delta(\omega + \omega_i + \omega_j) \right] \\
+ G_{ij}(\omega) \left[ \delta(\omega + \omega_i + \omega_j) + \delta(\omega + \omega_i - \omega_j) \right] \right] \\
+ f_\epsilon(\omega).
\]

(2.24)

The term indicator frequencies for the quadratic model are summarized in Table 2.1. For example, let \( p = 2 \) (two factors) and choose \( \omega_1 = .45, \omega_2 = .31 \). The resulting term indicator frequencies are as follows: (linear) .31, .45, (quadratic) .38, .10, (interaction) .24, .14. Observe the effect of aliasing on the frequencies. A term indicator frequency \( \nu \) which falls between .5 and 1 is observed at the frequency \( 1 - \nu \) (see, e.g., Priestley; 1981, p. 224). Thus, while \( 2\omega_1 = .62 \), the observed frequency is at \( 1 -.62 = .38 \). Similarly, \( \omega_1 + \omega_2 = .76 \), but the corresponding term indicator frequency is observed at \( 1 -.76 = .24 \).

A question immediately arises: can the driving frequencies \( \omega_1, \ldots, \omega_n \) be selected so that the term indicator frequencies correspond to a unique term in the proposed DRSW? The above indicates that in general the answer may be "no" for all term indicator frequencies, but that it is possible for those
corresponding to the highest degree of each factor (see Jacobson, Buss, and Schruben, 1987). Thus, in the quadratic model above the term indicator frequencies corresponding to linear terms are confounded by second order terms, but the quadratic and interaction frequencies are unconfounded. Observe that any factor oscillated about zero \( a_j = 0 \) will not confound lower order terms in the quadratic model, but that, e.g. cubic terms will always confound linear terms, even when oscillated about zero. If \( a_j \neq 0 \), then the quadratic terms will be confounded as well. For our analysis it is therefore most desirable to parameterize the model so that \( a_j = 0 \). In section 2.2 an algorithm is given for determining driving frequencies.

One potential difficulty involves gain. One way in which gain may affect the DRSM is the appearance of the noise spectrum \( f_{\xi}(\omega) \) in the response spectrum \( f_{\gamma}(\omega) \) (Equation 2.26). If we assume that the noise \( \xi(t) \) is "white", then then its spectrum \( f_{\xi}(\omega) \) is constant and will not affect the location of peaks. However, in general, peaks in the output spectrum may be due to in \( f_{\xi}(\omega) \). Schruben and Cogliano (1987) propose a independent second control run in which parameters are held constant. The power spectrum \( f_{\gamma}(\omega) \) from this run gives an independent estimate of the noise spectrum. The presence of peaks in the signal run, in which parameters are oscillated, may be detected by taking the ratio \( f_{\gamma}(\omega)/f_{\gamma}(\omega) \). This is analogous to the "signal to noise ratio" of signal processing. Using the asymptotic distribution of periodogram estimators, they construct an F-test based on the signal to noise ratio to detect the presence of peaks.
<table>
<thead>
<tr>
<th>Term</th>
<th>Indicator Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$\omega_i$ $i = 1, \ldots, n$</td>
</tr>
<tr>
<td>Quadratic</td>
<td>$2\omega_i$ $i = 1, \ldots, n$</td>
</tr>
<tr>
<td>Interaction</td>
<td>$\omega_i \pm \omega_j$ $i \neq j$</td>
</tr>
</tbody>
</table>

Table 2.1. Term Indicator Frequencies for the Quadratic Model
2.2. Selection of Driving Frequencies

Driving frequencies should be chosen to avoid confounding term indicator frequencies. Confounding occurs when term indicator frequencies corresponding to distinct terms are identical. For example, consider a quadratic model with two (or more) factors with driving frequencies of $\omega_1 = .2$ and $\omega_2 = .4$, corresponding to factors $x_1$ and $x_2$. The $x_1^2$ term is confounded with the $x_2$ term, since $2\omega_1 = \omega_2$: a peak in the output spectrum at .4 could be from a linear term in $x_2$, a second order term in $x_1$, or both.

A second consideration for selecting the driving frequencies is that of maximizing the smallest spacing between two distinct term indicator frequencies. This spacing determines the estimator bandwidth required to distinguish term indicator frequencies. In the example in Section 2.1 the minimum spacing is .04, the distance between term indicator frequencies .10 and .14. Ideally, all term indicator frequencies would be evenly spaced. Apparently, this can only occur in the case of purely linear terms (i.e. a linear model) or a quadratic model with only two factors. We assume initially that all potential terms are present in the model. After an initial factor screening experiment the experimenter may decide that certain terms are not present. The experimenter may, in this case, be willing to have confounding at corresponding term indicator frequencies for future experiments on the same model. We shall discuss this later.

A mixed integer linear program is given in Cogliano (1982) which obtains the optimal (maximal) minimum spacing for a model with given order and number of factors:

$$\max \min |\eta_i - \eta_j|$$
where the minimum is taken over all distinct term indicator frequencies $\eta_i$ and $\eta_j$ corresponding to given driving frequencies, and the maximum is taken over all possible selections of driving frequencies. This program is extremely difficult to solve for more than a few factors and a second order (or higher) model. However, once found, they can be tabulated for future use. Table 2.3 contains the best driving frequencies found to date for the second order model. A heuristic of Jacobson, Buss, and Schruben (1987) obtains a fairly good set of driving frequencies and occasionally gives optimal ones. We now briefly describe this heuristic.

Since the optimal frequencies are the solutions to the above mixed integer linear program having all integer coefficients, the frequencies must be rational numbers. Therefore, the frequency selection problem may be decomposed into two stages:

1) Selection of numerators.

2) Selection of the common denominator.

We shall present the heuristic for the second order model only; it is easily generalized to higher order models. The strategy in stage 1 is to start with a positive integer $p_1$, which will be the numerator for the first driving frequency, and eliminate $2p_1$ (corresponding to the quadratic term associated with the first frequency) from further consideration. The smallest candidate for the second numerator is $p_1$ plus the smallest integer not eliminated; call this number $\tilde{p}_2$. Now compute $2\tilde{p}_2$ and $\tilde{p}_2 \pm p_1$ (corresponding to quadratic and interaction terms, respectively). If these are all distinct, then $\tilde{p}_2$ is accepted as a numerator and now called simply $p_2$. The numbers $2p_2$ and $p_2 \pm p_1$ are eliminated from further consideration. If there is confounding with $2\tilde{p}_2$ and $\tilde{p}_2 \pm p_1$, then $\tilde{p}_2$ is replaced by $p_1$ plus the next largest integer not eliminated for $p_1$. This is continued until a
numerator has been selected for each factor.

To illustrate stage 1, let us find numerators for a second order model with three factors. Start with $p_1 = 1$ so that $\{1,2\}$ are eliminated from further consideration. The smallest candidate is now $1 + 3 = 4$, since 3 is the smallest integer not eliminated. This numerator adds 8 (quadratic), 3, and 5 (interaction) to the list of term numerators, which is now $\{1,2,3,4,5,8\}$. Finally, the smallest candidate for the third numerator is $4 + 6 = 10$, since 6 is the smallest number not eliminated. The term numerators added by 10 are $\{10,20,6,14,9,11\}$, so the final list is $\{1,2,3,4,5,6,8,9,10,11,14,20\}$ (see Table 2.2). The driving numerators are $\{1,4,10\}$.

Stage 2 finds the best (smallest) denominator corresponding to the numerators obtained in stage 1. Clearly a feasible denominator can be obtained by adding 1 to twice the largest term numerator. However, this is usually not best for a given set of numerators. Stage 2 attempts to exploit the aliasing effect to reduce the size of the denominator. For a second order model, the largest numerator corresponds to the quadratic term of the largest driving numerator. A small denominator will result if this term can be aliased to a small number. This corresponds to finding a point at which to "fold" the numerators so that all term numerators larger than the fold point fall onto vacant numbers (i.e. onto integers which are not term numerators). To illustrate stage 2 for the example above, first try to alias 20 to 7 (see figure 2.1). This corresponds to a fold at $13\%$, so that 14 is also aliased to 13. The denominator corresponding to this is $13\% \times 2 = 27$, and the minimum spacing is only $1/54$ (This is the distance from $13/27$ to $1/2$). The next integer to which we can alias 20 is 12, resulting in a fold point of 16 and a denominator of 32 (see Figure 2.1). Thus, the driving
frequencies are \{1/32,4/32,10/32\} and the term indicator frequencies are
denominator corresponding to optimal numerators \{1,4,11\} is 28, obtained by
the fold point of 14 (see Figure 2.2).

If the heuristic starts with a numerator other than 1, then there are
other possibilities. For example, the best fold for a given set of term
numerators need not be at an integer. If the heuristic starts with numerator
2, then driving numerators of \{2,3,10\} result, with term numerators as in
Figure 2.3. The best fold for these term numerators is at 14\% and the
denominator is 29, only slightly worse than the optimal denominator of 28.
Some starting values require no fold: if the heuristic starts at 4, then
\{4,5,7\} are the driving numerators, 14 is the largest term numerator, and the
denominator is 30 (see Figure 2.4).

Table 2.3 gives the optimal frequency selections for a second order
model with up to 7 factors. The heuristic may be modified to allow skipping
eligible driving numerators. For example, if the second eligible numerator
is chosen at each step, then (starting at 1) the resulting driving numerators
are \{1,5,8\}. Folding at 14 yields a denominator of 28, the optimal
denominator (see Figure 2.5). Observe that optimal (evenly spaced)
umerators result if the third numerator is chosen to be 11 instead of 10,
resulting in term numerators \{1,2,3,4,5,7,8,10,11,12,15,22\}.

Table 2.3 gives sets of frequencies for various numbers of factors with
a second order model. The frequencies for 2-7 factors are optimal, assuming
a full model (all terms present). The frequencies for 8-12 factors are not
necessarily optimal, but are the best available at present. The table gives
numerators and denominators separately, and for the factors with optimal
spacing all the alternative frequencies are given. For example, if there are
5 factors, the experimenter may choose driving frequencies
\{1/69, 4/69, 13/69, 19/69, 29/69\}, \{4/69, 5/69, 7/69, 20/69, 26/69\}, or
\{2/69, 5/69, 11/69, 25/69, 26/69\}. These are the only ones with optimal spacing;
the choice of any other driving frequencies results in a narrower minimum
spacing. See Jacobson, Schruben, and Buss (1986) for further details.

If the experimenter has decided that certain terms are not in the model,
then better frequency selection (i.e. larger minimum spacing) usually
results. Typically, models encountered in practice are not "full" in the
sense that not all possible terms are present. In such cases, a larger
minimum spacing results from there being fewer term indicator frequencies.

For example, consider a second order model with three factors. Suppose
the experimenter has decided, perhaps on the basis of a previous frequency
domain factor screening experiment, that the only terms in the model are
linear, together with \(x_1^2\), \(x_1x_2\), and \(x_1x_3\). The above heuristic may be applied
with the modification that only the term indicator numerators for these terms
are considered. Taking 1 to be the numerator for \(x_1\), the \(x_1^2\) term gives term
numerator of 2. The smallest numerator for \(x_2\) is therefore 4, with term
numerator for \(x_1x_2\) 3 and 5. Finally, 7 may be chosen as the numerator for
\(x_3\) with term numerators for \(x_1x_3\) of 6 and 8. Thus, the term numerators are
\{1, 2, 3, 4, 5, 6, 7, 8\}, which are equally spaced. The denominator is therefore 18
with minimum spacing of 1/18, a considerable improvement over the optimal
minimum spacing for the full model, 1/28. If \(x_2x_3\) were in the model as well,
then 10 would have to be chosen for the \(x_3\) numerator. The resulting term
numerator \{1, 2, 3, 4, 5, 6, 9, 10, 11, 14\} have a denominator of 26 and minimum
spacing 1/26, produced by folding at 13. Apparently the optimal driving
numerator in this case are \{1, 5, 8\} with term numerators of
\{1, 2, 3, 4, 5, 6, 7, 8, 9, 13\}, denominator 23, with a fold at 11%. If \(x_2^2\) is also
present (so that the model is full except for $x_{32}$), the same driving numerators give term numerators of $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13\}$ with a denominator of 24 and a fold at 12. These are equally spaced, and therefore optimal.
Table 2.2 Heuristic for Frequency Selection.

First Numerator = 1

<table>
<thead>
<tr>
<th>Step</th>
<th>Linear</th>
<th>Quadratic</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>8</td>
<td>3 5</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>20</td>
<td>6 9 11 14</td>
</tr>
</tbody>
</table>
Table 2.3 Frequencies for Various Numbers of Factors,
Second Order Model. Frequencies for 2–7 Factors
are Optimally Spaced.

<table>
<thead>
<tr>
<th># Factors</th>
<th>Denominator</th>
<th>Numerators</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>14</td>
<td>(1,4), (2,3)</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>(1,5,8), (1,4,11), (3,4,13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3,5,12), (8,9,13), (9,11,12)</td>
</tr>
<tr>
<td>4</td>
<td>46</td>
<td>(1,4,10,17), (6,8,9,13), (2,3,11,18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4,5,7,20), (3,5,12,16), (2,9,10,15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1,6,16,19)</td>
</tr>
<tr>
<td>5</td>
<td>69</td>
<td>(1,4,13,19,29), (4,5,7,20,26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2,5,11,25,26)</td>
</tr>
<tr>
<td>6</td>
<td>103</td>
<td>(1,11,28,31,35,49), (3,4,13,28,40,42)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4,9,10,21,37,44), (6,32,40,42,43,47)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(8,15,18,20,29,42), (10,12,15,21,28,29)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11,27,35,36,48,50), (16,23,24,43,45,49)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(20,24,30,42,45,47)</td>
</tr>
<tr>
<td>7</td>
<td>130</td>
<td>(1,4,19,31,44,53,60), (4,7,9,24,30,49,59)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1,9,14,21,40,46,57), (1,7,18,22,27,57,60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2,3,12,29,37,50,57), (7,9,17,20,32,62,63)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3,7,19,28,30,43,48), (1,10,16,29,34,37,41)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3,8,10,27,41,42,63), (2,17,22,23,30,33,59)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6,9,19,20,36,41,43), (6,31,40,47,51,61,64)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3,21,41,49,50,54,64), (10,11,36,44,49,51,63)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(8,33,47,48,51,53,60), (11,16,17,20,46,59,61)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(10,11,18,23,37,53,62), (19,29,30,33,54,56,61)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11,14,23,24,29,31,50), (21,27,34,51,56,59,60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(12,17,21,27,40,47,58), (31,32,38,40,43,53,57)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(20,33,37,42,43,58,61), (23,28,38,49,49,50,63)</td>
</tr>
<tr>
<td>8</td>
<td>168</td>
<td>(10,16,29,33,38,40,41,75)</td>
</tr>
<tr>
<td>9</td>
<td>209</td>
<td>(8,30,33,39,40,44,57,59,94)</td>
</tr>
<tr>
<td>10</td>
<td>268</td>
<td>(10,12,13,27,31,59,65,94,101,110)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2,15,22,33,47,56,75,83,121,126)</td>
</tr>
<tr>
<td>11</td>
<td>340</td>
<td>(19,21,22,26,32,46,55,105,117,135,166)</td>
</tr>
<tr>
<td>12</td>
<td>448</td>
<td>(27,44,56,57,59,63,77,102,124,150,155,219)</td>
</tr>
</tbody>
</table>
Figure 2.1 Denominator = 27 by "Folding" at 14\%
Bandwidth = \(0.5 - \frac{13}{27} = \frac{1}{54}\)

Figure 2.2 Denominator = 28 by Folding at 16.
Bandwidth = \(\frac{1}{28}\)

Figure 2.3 Starting Heuristic with Numerator = 2
Denominator = 29 by Folding at 14\%
Bandwidth = \(\frac{1}{29}\)
Figure 2.4 Starting Heuristic with Numerator = 4
No Folding is Required
Denominator = 30, Bandwidth = 1/30.

Figure 2.5 Driving Numerators = \{1, 5, 8\}
Denominator = 28 by Folding at 14
Bandwidth = 1/28 (Optimal).
2.3. **Periodic Steady State**

Consider a frequency domain experiment in which there is at least one factor present. As demonstrated above, this may be represented as a DRSM in which the \( x(t) \) terms are products of cosines. Clearly such a system cannot reach steady state as it is usually defined in the usual way for simulations or stochastic processes. However, under the assumptions for frequency domain experiments, asymptotically the output reaches a situation in which the covariance is stationary and the mean oscillates in a deterministic fashion. This is similar to the concept of a *periodic attractor* in dynamical systems (Hirsh and Smale, 1974) in which the orbits (solutions) of systems are periodic. Thus, we say that a simulation has *periodic steady state* if there is a deterministic periodic function \( \mu \) such that \( y(t) - \mu(t) \) is stationary. We can choose \( \mu \) without loss of generality to be the mean of \( y(t) \).

To clarify this idea, consider a child on a swing. If there is no movement then the swing and the child eventually come to rest. If the child starts pumping, however, then eventually the child and the swing will oscillate in a more or less consistent manner, depending on the nature of the swing and the forcefullness of the pumping. As long as the pumping continues, the motion will remain consistent. It seems reasonable to call this motion "steady state" of the system with the periodic input (pumping).

No consider a system described by a (static) response surface model and the resulting system obtained by performing a frequency domain experiment on that system. Suppose the result of the frequency domain experiment may be represented as a DRSM in which the noise process \( \{ \varepsilon(t) \} \) reaches stationarity (in the sense of stochastic processes). Then the output will reach periodic steady state, and the resulting spectral analysis to determine significant factors will be valid.
2.5. *Synchronous vs Asynchronous Simulations*

Discrete-event simulations may be classified according to how they handle the passage of time (Bratley, Fox, and Schrage, 1987). Synchronous (or time-slice) simulations advance the simulation clock by a fixed amount $\Delta t$. Asynchronous simulations advance the clock only when modeled events occur. Despite the disadvantages to implementing synchronous simulations, they are frequently used in applications, being *conceptually* simpler. One advantage of synchronous simulations is that frequency domain experiments are easier to implement. At each time $k\Delta t$, the value of factor $j$ is simply $a_j + \cos 2\pi(k\Delta t)\omega_j$. The logic during that time step is identical to the non-oscillated case. Furthermore, there are no difficulties in interpreting the time mechanism, since the simulation clock has been modeled to increment with "real" time.

On the other hand, asynchronous simulations present some difficulties in the direct implementation of frequency domain experiments. For such models, the clock is incremented at arbitrary simulation times, and nothing happens in the model during the intervening time intervals. Thus, an asynchronous simulation amounts to a random scaling of ("real") time. This affects frequency domain experiments as follows. If a parameter is chosen to be oscillated at a high frequency using simulation time, then it will typically cycle many times during the time intervals between occurrences of events. The resulting frequency that the simulation "sees" will be much lower.

To get around this difficulty for a queueing model, Schruben and Cogliano (1987) use the customer number as the clock with respect to which the parameters are oscillated. This approach may be used in situations for which the customers depart in the same order they arrive. It may even be used for cases in which the customers depart in different order; on output,
they are simply re-sorted to their original sequence (see Sargent, Tom, and Schruben, 1987). However, for example, for queueing systems with more than one customer type, this approach does not seem to work. One suggestion has been to use the (global) simulation clock anyway, but even that appears to present difficulties (Jacobson, Morrice, and Schruben, 1988).

Thus, the simulation experimenter who is using a synchronous simulation model should have no difficulty implementing frequency domain experiments. The simulation clock and the oscillatory clock are identical. On the other hand, the experimenter who is using an asynchronous simulation model must use more caution and design the frequency domain experiment more carefully. It may be necessary to run different experiments for each set of factors which may be synchronized. That is, the set of factors is grouped into sets such that if the factors in a given set are the only ones oscillated, then a well-defined oscillatory clock may be defined. Signal and noise runs are performed on each of these sets. The final step is a fractional factorial experiment involving only those factors determined to be influential. This final experiment should be designed to identify possible interactions between factors in different sets. Of course, factors in multiple sets would have several "chances" to be influential, and there could be conflicting outcomes for the same factor when run in different sets. This would be evidence of either model mis-specification or some kind of factor-frequency interaction.
3. Example of Frequency Domain Experiments for Factor Screening

In this section we will discuss some frequency domain experiments for factor screening. The model is a large-scale finite element model of geological processes related to the region surrounding a hypothetical nuclear waste repository, called the Geological Simulation Model (GSM) (Petrie, et al, 1986). This model is in the development stage at Battelle Pacific Northwest Laboratories as one component of the process of site selection.

3.1. The Geological Simulation Model

GSM is a large scale finite element model with randomness included as part of the geologic description (Petrie, et al [1986]). Characteristics of the rock and soil in the region are specified at points of the finite element grid, and at each time step random processes such as climate change and tectonics influence the system behavior. The primary output variable of interest is the groundwater flow at a particular location (the hypothetical repository site). The location of such a repository will, in part, be determined by the geologic stability of the region and by the relative absence of groundwater flow, since any release of contents which reaches groundwater will have serious environmental consequences.

One issue that GSM attempts to address is the relative importance of the various geological processes and of the specific values of certain characteristics of the model. This makes GSM a good candidate for frequency domain methodology. The processes that are emphasized in the frequency domain experiments are recharge, slope retreat, and denudation (erosion).

Recharge refers to the replenishment of the groundwater by precipitation. GSM treats recharge directly, rather than having separate variables for precipitation, temperature, vegetation, and other variables
which affect recharge. At each time step the total recharge for the region is generated from a probability distribution. This amount is distributed about the region according to a deterministic weighting function. This approach is followed for a number of variables in GSM.

The second process considered is slope retreat. Slope retreat is the lateral movement of the earth. Its impact on the model is to change the location of the finite element nodes.

The third process considered is denudation, or erosion. The effect of erosion is to lower the elevation of surface nodes of the grid. Denudation is represented in a similar manner as recharge: a single random variable is generated representing the total denudation for the region. The relative values are then determined according to a deterministic weight function. Following generation, the affected grid nodes are updated.

First, a noise run of 128 time steps was performed; recall that this involved running the model with the parameters held fixed at their nominal values. The periodogram estimate for this run is shown in Figure 3.1, and demonstrates a rather flat spectrum. Following Schruben and Cogliano (1987), common random number seeds were used for the signal runs. The first signal experiment on GSM involved all three variables. The model was run for 512 time steps with variables oscillated according to the frequencies shown in Table 3.1, with the corresponding term indicator frequencies also shown in Table 3.1. The spectrum for this run is shown in Figure 3.2. The large peak at frequency .393 corresponds to a linear term in total recharge. The absence of peaks at other frequencies indicates a lack of dependence on those variables. To further confirm this, two more signal runs were made: one oscillated just denudation and slope retreat, and the other only recharge. The resulting spectrum for denudation/slope retreat shown in Figure 3.3 is
indistinguishable from the noise run, while the spectrum for recharge only is shown in Figure 3.4 and confirms the strong linear response.

Observe that for the last two experiments that the variables were oscillated at frequencies which were different than those of the first signal run. The consistency of results together with the relatively flat spectra in runs 1 and 3 indicate that gain is probably not a factor here. Apparently recharge is the most influential parameter in the model, while the other parameters have a negligible effect on the output.

Another experiment focused on recharge and the way in which it was modelled in GSM. At each time step the total recharge for the region is generated. The recharge for each node in the finite element grid is determined by a relative weight function (rwf), which is piecewise linear over the cross section of the region (see Petrie, et. al., 1986). First, the rwf was taken to be linear across the region, and therefore only determined by its endpoints. The height of the endpoints was oscillated between 0 and 1 while the total recharge was oscillated as well. The spectrum from this experiment is shown in Figure 3.5. The peaks at .036 and at .0143 indicate that total recharge affect the output linearly, while the absence of a peak at .393 or at .214 indicate that the right weight does not influence the output. There are small but noticeable peaks at .107 and .179, which indicate an interaction between total recharge and the left rwf weight.

Several points need to be made. The first is that the absence of effect by the variables may be more important information than the presence of effect. This information can be useful in several ways. If the model is under development (as is GSM) variables which do not affect the output are less important to model more precisely than the ones which are shown to affect the output significantly. Thus, modelling effort and data collection
may be oriented towards those parameters which prove to be influential. Conversely, if knowledge about the real system says that certain variables are in fact influential, then the failure of the model to reflect this influence in its behavior is cause to focus on those variables to make the model more realistic and improve model validity.
Table 3.1. Term Indicator Frequencies: Three Parameters, Second Order

<table>
<thead>
<tr>
<th>Terms</th>
<th>Indicator Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>.036</td>
</tr>
<tr>
<td></td>
<td>.143</td>
</tr>
<tr>
<td></td>
<td>.393</td>
</tr>
<tr>
<td>Quadratic</td>
<td>.072</td>
</tr>
<tr>
<td></td>
<td>.286</td>
</tr>
<tr>
<td></td>
<td>.214</td>
</tr>
<tr>
<td>Interaction</td>
<td>.179</td>
</tr>
<tr>
<td></td>
<td>.464</td>
</tr>
<tr>
<td></td>
<td>.329</td>
</tr>
<tr>
<td></td>
<td>.107</td>
</tr>
<tr>
<td></td>
<td>.367</td>
</tr>
<tr>
<td></td>
<td>.250</td>
</tr>
</tbody>
</table>
Figure 3.1. Frequency Response
"Noise Run"
0.000  0.742e-01+
0.010  1.187e-01+
0.020  3.045e-03+
0.030  3.653e-03+
0.040  4.951e-03+
0.050  4.796e-03+
0.060  2.272e-03+
0.070  3.791e-03+
0.080  6.874e-03+
0.090  7.119e-03+
0.100  2.848e-03+
0.110  5.203e-03+
0.120  2.949e-03+
0.130  5.049e-03+
0.140  3.348e-03+
0.150  6.469e-03+
0.160  6.820e-03+
0.170  3.743e-03+
0.180  5.290e-03+
0.190  2.761e-03+
0.200  5.402e-03+
0.210  5.867e-03+
0.220  5.182e-03+
0.230  3.159e-03+
0.240  1.427e-03+
0.250  3.014e-03+
0.260  5.543e-03+
0.270  3.563e-03+
0.280  4.947e-03+
0.290  6.108e-03+
0.300  0.786e-02+
0.310  4.617e-03+
0.320  3.600e-03+
0.330  3.327e-03+
0.340  2.693e-03+
0.350  -6.615e-03+
0.360  -1.336e-02+
0.370  -4.945e-02+
0.380  -3.169e-01+
0.390  1.287e+01+*****************************************************
0.400  3.698e+00+**********
0.410  1.404e-01+
0.420  3.026e-02+
0.430  1.301e-02+
0.440  0.934e-02+
0.450  7.469e-03+
0.460  5.072e-03+
0.470  7.280e-03+
0.480  1.601e-03+
0.490  3.760e-03+
0.500  2.087e-03+

Figure 3.2. Frequency Response

Total Recharge: \( \omega_1 = 0.393 \)

Denudation: \( \omega_2 = 0.143 \)

Slope Retreat: \( \omega_3 = 0.036 \)
Figure 3.3: Frequency Response

Denudation: $\omega_1 = 0.071$

Slope Retreat: $\omega_2 = 0.286$
\begin{tabular}{lrr}
0.000 & 0.966e-01+ \\
0.010 & 1.175e-01+ \\
0.020 & 0.627e-01+ \\
0.030 & 4.126e+00+ \textbf{*} \\
0.040 & 7.678e+00+ \textbf{*} \\
0.050 & -1.121e-01+ \\
0.060 & -1.877e-02+ \\
0.070 & -4.351e-03+ \\
0.080 & -6.080e-03+ \\
0.090 & 1.650e-04+ \\
0.100 & 1.706e-01+* \\
0.110 & 6.077e-01+** \\
0.120 & -4.622e-02+ \\
0.130 & -1.470e-01+ \\
0.140 & 5.443e+00+ \textbf{*} \\
0.150 & 1.601e+00+***** \\
0.160 & 0.714e-01+ \\
0.170 & 5.231e-02+ \\
0.180 & 8.156e-01+*** \\
0.190 & -1.891e-03+ \\
0.200 & 0.829e-02+ \\
0.210 & 5.195e-03+ \\
0.220 & 7.750e-03+ \\
0.230 & 4.278e-03+ \\
0.240 & 3.813e-03+ \\
0.250 & 3.840e-03+ \\
0.260 & 3.541e-03+ \\
0.270 & 3.633e-03+ \\
0.280 & 1.121e-02+ \\
0.290 & 2.687e-02+ \\
0.300 & 4.377e-03+ \\
0.310 & 1.818e-03+ \\
0.320 & 3.474e-03+ \\
0.330 & 4.101e-03+ \\
0.340 & 3.616e-03+ \\
0.350 & 2.502e-03+ \\
0.360 & 2.543e-03+ \\
0.370 & 3.559e-03+ \\
0.380 & 3.279e-03+ \\
0.390 & 2.742e-03+ \\
0.400 & 3.093e-03+ \\
0.410 & 2.939e-03+ \\
0.420 & 3.217e-03+ \\
0.430 & 1.962e-03+ \\
0.440 & 1.996e-03+ \\
0.450 & 2.466e-03+ \\
0.460 & 3.389e-03+ \\
0.470 & 1.090e-03+ \\
0.480 & 1.940e-03+ \\
0.490 & 1.954e-03+ \\
0.500 & 3.011e-03+ \\
\end{tabular}

\textbf{Figure 3.4. Frequency Response}

Total Recharge: $\omega_1 = .036$
Left RWF: $\omega_2 = .147$
Right RWF: $\omega_3 = .343$
4. Estimating the DC Response

In this section we address the relationship between frequency domain and conventional simulation experiments. In the frequency domain experiments described in preceding sections, inputs which normally would be held constant are oscillated. In the electrical metaphor, the frequency analysis of the output is analogous to obtaining the AC response of the system at specific frequencies. On the other hand, a simulation performed in the conventional manner (which purports to mimic the actual conditions of the simulated system) has the inputs held constant. In the electrical metaphor, this output is the "DC response". In the DRSM this is the same as setting \( x(t) \equiv x \) for all \( t \). For each pseudo-linear term \( x^\alpha(t) \) in the model, the corresponding response is therefore

\[
x^\alpha \int_0^\infty g_\alpha(\tau) \, d\tau.
\]  
(4.1)

Observe that the integral in expression (4.1) is the Fourier transform of \( g_\alpha \) evaluated at zero \( \hat{g}_\alpha(0) \).

4.1. The Kramers-Kronig Relations

It turns out that a pair of identities (known as the Kramers-Kronig relations in the electrical engineering literature (see Jackson [1975] p. 310) relate the real and imaginary parts of the Fourier transform of impulse response functions for physically realizable systems. These are:

\[
\text{Re} \hat{g}(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\text{Im} \hat{g}(\eta)}{\eta - \omega} \, d\eta
\]  
(4.2)

\[
\text{Im} \hat{g}(\omega) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\text{Re} \hat{g}(\eta)}{\eta - \omega} \, d\eta.
\]  
(4.3)
The integrals in Equations (4.2-3) are taken in the Cauchy principal value sense (see, e.g., Stein & Weiss, 1971). For example, the integral in Equation (4.2) is defined as

$$\lim_{\delta \to 0} \left[ \int_{-\infty}^{\omega-\delta} + \int_{\omega+\delta}^{\infty} \frac{\mathfrak{m} \hat{g}(\eta)}{\eta - \omega} \, d\eta \right]. \tag{4.4}$$

Let us illustrate the Kramers-Kronig relations in Equations (4.2-3) with a simple exponential function $g(t) = e^{-\gamma t}$. The corresponding Fourier transform is $\hat{g}(\omega) = 1/(\gamma - i \omega) = (\gamma + i \omega)/(\gamma^2 + \omega^2)$. Now,

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\mathfrak{m} \hat{g}(\eta) \, d\eta}{\eta - \omega} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\eta \, d\eta}{(\eta - \omega)(\gamma^2 + \eta^2)}$$

(Expanding the integrand in partial fractions)

$$= \frac{1}{\pi} \cdot \frac{1}{\gamma^2 + \omega^2} \int_{-\infty}^{\infty} \left[ \frac{\omega}{\eta - \omega} - \frac{\omega \eta}{\gamma^2 + \eta^2} + \frac{\gamma^2}{\gamma^2 + \eta^2} \right] \, d\eta$$

(The first two terms are zero in the principle value sense)

$$= \frac{1}{\pi} \cdot \frac{\gamma}{\gamma^2 + \omega^2} \int_{-\infty}^{\infty} \frac{d\eta}{\gamma^2 + \eta^2}$$

$$= \frac{\gamma}{\gamma^2 + \omega^2} \Rightarrow \Re \hat{g}(\omega).$$

Similarly, for (4.3),

$$- \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\Re \hat{g}(\eta) \, d\eta}{\eta - \omega} = - \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\gamma \, d\eta}{(\eta - \omega)(\gamma^2 + \eta^2)}$$
\[
\begin{align*}
&= -\frac{\gamma}{\pi} \cdot \frac{1}{\gamma^2 + \omega^2} \int_{-\infty}^{\infty} \left[ \frac{1}{\eta - \omega} - \frac{\eta}{\gamma^2 + \eta^2} - \frac{\omega}{\gamma^2 + \eta^2} \right] \, d\eta \\
&= \frac{\gamma \omega}{\pi} \cdot \frac{1}{\gamma^2 + \omega^2} \int_{-\infty}^{\infty} \frac{d\eta}{\gamma^2 + \eta^2} \\
&= \frac{\omega}{\gamma^2 + \omega^2} = \Im \hat{g}(\omega).
\end{align*}
\]

We can now use the Kramers-Kronig relations to approximate the "DC response" for the system of Equation (4.1); that is, we can approximate \( \hat{g}(0) \). Since \( \hat{g}(0) \) is real, Equation (4.2) becomes, for \( \omega = 0 \)
\[
\hat{g}(0) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\Im \hat{g}(\eta)}{\eta} \, d\eta 
\]
(4.5)
or, since \( \Im \hat{g}(\eta)/\eta \) is an even function,
\[
\hat{g}(0) = \frac{2}{\pi} \int_{0}^{\infty} \frac{\Im \hat{g}(\eta)}{\eta} \, d\eta.
\]
(4.6)

We now approximate the integral in Equation (4.6) with a Riemann sum. Choose a small \( \omega_0 \) (the smaller the better) and evaluate the integrand at points \( k\omega_0 \), for \( k \) an odd integer, over intervals \([ (k-1)\omega_0, (k+1)\omega_0 ] \). The approximation is
\[
\hat{g}(0) \approx \frac{2}{\pi} \sum_{k=1 \text{ odd}}^{\infty} \frac{\Im \hat{g}(k\omega_0)}{k\omega_0} \cdot 2\omega_0
\]
\[
= \frac{4}{\pi} \sum_{k=1 \text{ odd}}^{\infty} \frac{\Im \hat{g}(k\omega_0)}{k}.
\]
(4.7)
To evaluate $\hat{g}(k\omega_0)$ for $k$ odd we put in a signal $\tilde{x}(t)$ with frequencies at $k\omega_0$ and analyze the spectrum of the output. One such signal is particularly simple: a square wave with period $T = 2\pi/\omega_0$ and height $1/2$, which has Fourier series

$$\tilde{x}(t) = \pi \sum_{k=1}^{\infty} \frac{\sin k\omega_0 t}{k}.$$  \hspace{1cm} (4.8)

The jumps of the square wave at $nT$ are of size 1 for $n$ even and -1 for $n$ odd. Let $\tilde{y}(t)$ be the observed output from the system with $\tilde{x}(t)$ as input. Then:

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

$$= \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} \int_0^\infty \sin k\omega_0 (t-\tau) g(\tau) \, d\tau$$

$$= \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} \sin k\omega_0 t \int_0^\infty \cos k\omega_0 \tau g(\tau) \, d\tau$$

$$- \frac{1}{k} \cos k\omega_0 t \int_0^\infty \sin k\omega_0 \tau g(\tau) \, d\tau$$

$$= \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} \sin k\omega_0 t \Re \hat{g}(k\omega_0) - \frac{1}{k} \cos k\omega_0 t \Im \hat{g}(k\omega_0).$$ \hspace{1cm} (4.9)

Thus, $-2\Im \hat{g}(k\omega_0)/k\pi$ is the Fourier coefficient of the $\cos k\omega_0 t$ term in the Fourier series expansion of $\tilde{y}(t)$. Therefore,

$$- \frac{2}{\pi k} \Im \hat{g}(k\omega_0) = \frac{2}{T} \int_0^T \cos k\omega_0 t \tilde{y}(t) \, dt$$ \hspace{1cm} (4.10)

so
\[ \hat{g}(k\omega_0) = -\frac{\pi k}{T} \int_0^\infty \cos k\omega_0 t \tilde{y}(t) \, dt. \]  \hspace{1cm} (4.11) 

Substituting Equation (4.11) into Equation (4.7), we obtain

\[
\hat{g}(0) = -\frac{4}{T} \sum_{k_{0\text{ odd}}}^\infty \int_0^T \cos k\omega_0 t \tilde{y}(t) \, dt
\]

\[
= -\frac{4}{T} \int_0^T \tilde{y}(t) \sum_{k_{0\text{ odd}}}^\infty \cos k\omega_0 t \, dt. \]  \hspace{1cm} (4.12)

Now, the sum in Equation (4.12), while not defined in the ordinary sense, is defined in the generalized function sense and is the sum of \(\delta\)-functions (see Appendix I). Indeed,

\[
\sum_{k_{0\text{ odd}}}^\infty \cos k\omega_0 t = \frac{\pi}{2\omega_0} \sum_{n=-\infty}^{\infty} [\delta(t-nT) - \delta(t-(n-\frac{1}{2})T)]. \]  \hspace{1cm} (4.13)

Putting Equation (4.13) into Equation (4.12), we have

\[
\hat{g}(0) = -\frac{2\pi}{T\omega_0} \int_0^T \sum_{n=-\infty}^{\infty} [\delta(t-nT) - \delta(t-(n-\frac{1}{2})T)] \tilde{y}(t) \, dt
\]

\[
= \tilde{y}(T/2) - \tilde{y}(T) . \]  \hspace{1cm} (4.14)

We could have integrated over \([0,NT]\) for some integer \(N\) instead of over \([0,T]\), since \(NT\) is also a period of \(\tilde{y}(t)\). In this case, Equation (4.14) becomes

\[
g(0) = \sum_{n=1}^N [\tilde{y}((n-\frac{1}{2})T) - \tilde{y}(nT)]. \]  \hspace{1cm} (4.15)
Let us examine the consequences of (4.14). The experiment consists of running the system with \( x(t) \) fixed at \( \frac{\pi}{2} \) for a period of \( T/2 \), then at \( -\frac{\pi}{2} \) for a period of \( T/2 \), etc. The approximation for \( \hat{g}(0) \), that is the coefficient of \( x \) in the static model, is the difference between the value of the output after having been run at \( \frac{\pi}{2} \) and the value of the output after having been run at \( -\frac{\pi}{2} \). If the system were truly linear and noiseless (as is implied by the assumptions so far) then this is reasonable so long as \( T/2 \) is sufficiently large so as to let the system nearly reach steady state. Observe that this is controlled by the choice of \( \omega_0 \). If the experiment is carried out for more cycles, the estimator is simply the average of the differences for each successive period, as Equation (4.15) indicates. Thus, the Kramers-Kronig approach has, for this simple case, an intuitively simple interpretation. The approach works, however, even when the interpretation is not so simple.

To illustrate, consider a model for which there are \( n \) linear terms:

\[
y(t) = \sum_{j=1}^{n} \int_{0}^{\infty} x_j(t - \tau) g_j(\tau) \, d\tau.
\]

(4.16)

Suppose that the inputs \( x_j(t) \) are written as a Fourier series containing only sine terms:

\[
x_j(t) = \sum_{k=1}^{\infty} b_{jk} \sin k\omega_j t
\]

(4.17)

so that the output is

\[
y(t) = \sum_{j=1}^{n} \int_{0}^{\infty} \sum_{k=1}^{\infty} b_{jk} \sin k\omega_j (t-\tau) g_j(\tau) \, d\tau.
\]
Expanding the sine by the sum of angles formula, we have

\[
y(t) = \sum_{k=1}^{\infty} \sum_{j=1}^{n} b_{jk} \left[ \sin k\omega_j t \int_0^\infty \cos k\omega_j \tau \, g_j(\tau) \, d\tau - \cos k\omega_j t \int_0^\infty \sin k\omega_j \tau \, g_j(\tau) \, d\tau \right].
\]

Recalling that \( \Re \hat{g}(\omega) = \int \cos(\omega \tau) g(\tau) \, d\tau \) and \( \Im \hat{g}(\omega) = \int \sin(\omega \tau) g(\tau) \, d\tau \), we have

\[
y(t) = \sum_{k=1}^{\infty} \sum_{j=1}^{n} b_{jk} \left[ \sin k\omega_j t \Re \hat{g}_j(k\omega_j) - \cos k\omega_j t \Im \hat{g}_j(k\omega_j) \right]. \tag{4.18}
\]

Thus, multiplying both sides of Equation (4.18) by \( \cos m\omega_j t \) and integrating, we have

\[
\frac{2}{T} \int_0^T \cos m\omega_j t \, y(t) \, dt = -\sum b_{jk} \Im \hat{g}_j(k\omega_j) \tag{4.19}
\]

where the sum in Equation (4.19) is taken over all \((j,k)\) such that \(k\omega_j = m\omega_i\).

Here the problem of choosing frequencies \(\omega_j\) is different from that of choosing frequencies for the factor screening experiments. In this case they should be chosen so that the right side of Equation (4.19) consists of a single term for each \((j,k)\). Ratios of relatively prime odd numbers is one solution. Thus, for finitely many \((j,k)\)'s, we can calculate \(\Im \hat{g}(k\omega_j)\) by:

\[
\Im \hat{g}_j(k\omega_j) = \frac{2}{b_{jk} T} \int_0^T \cos k\omega_j t \, y(t) \, dt \tag{4.20}
\]

where \(T\) is the common period of \(\omega_1, \ldots, \omega_n\).
It turns out that the frequency selection problem in this instance is identical with that of a similar situation in which the impulse response functions are assumed to have a particular form. One approach is given in Section 5.3.

4.2. Example: M/M/1 Queue

We will illustrate the use of the Kramers-Kronig relations for determining static (DC) response with an M/M/1 queueing example. We will consider two parameters: \( \lambda \) = the arrival rate and \( \mu \) = the service rate. If \( W_n \) is the waiting time of the \( n \)th customer, then the well known recursion holds:

\[
W_{n+1} = [W_n + v_n - u_n]^+ \tag{4.21}
\]

In Equation (4.21) \( \{u_n\} \) and \( \{v_n\} \) are the inter-arrival times and service times, each a sequence of independent identically distributed random exponential random variables with means \( 1/\lambda \) and \( 1/\mu \), respectively (see Feller [1979]). The mean waiting time in steady state is

\[
\bar{W} = \frac{1}{\mu - \lambda}. \tag{4.22}
\]

The objective here is to approximate the derivatives of \( \bar{W} \) with respect to \( \lambda \) and \( \mu \). These are

\[
\frac{\partial \bar{W}}{\partial \lambda} = \frac{1}{(\lambda - \mu)^2} \tag{4.23}
\]

and
\[
\frac{\partial \bar{W}}{\partial \mu} = \frac{-1}{(\lambda - \mu)^2}.
\]  

(4.24)

The parameters \( \lambda \) and \( \mu \) were varied as square waves throughout the run with amplitudes \( \Delta \lambda = \Delta \mu = .05 \), center points .8 and 1.0, and half-periods 5 and 5\( \frac{1}{2} \), respectively. For this example "time" was the customer number. For each period an approximation to the derivative was computed by taking the difference of the waiting time after a "high" period and the waiting time after a "low period" and dividing by twice the amplitude. Four runs of 3000 customers were performed, with the results shown in table 4.1. For this example, the true values are 25 and -25 for \( \partial \bar{W} / \partial \lambda \) and \( \partial \bar{W} / \partial \mu \), respectively. Observe that the procedure does get the sign and rough order of magnitude correct in each case, despite the fact that the procedure has only been justified in the case of one parameter without noise.

4.3. Discussion

Clearly this approach needs much further development to prove useful. So far, we have derived a procedure based upon a linear noiseless system; such systems are well-studied and understood. Our orientation has been towards non-linear systems with noise, and it is not readily apparent that the Kramers-Kronig relations have more to tell us about them. However, the limited success in the preceding queueing example is grounds for optimism.
<table>
<thead>
<tr>
<th>run</th>
<th>$\frac{\partial \bar{W}}{\partial \mu}$</th>
<th>$\frac{\partial \bar{W}}{\partial \lambda}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-31.8</td>
<td>22.6</td>
</tr>
<tr>
<td>2</td>
<td>-37.5</td>
<td>29.0</td>
</tr>
<tr>
<td>3</td>
<td>-39.1</td>
<td>30.9</td>
</tr>
</tbody>
</table>

Table 4.1: Kramers-Kronig Procedure Applied to M/M/1 Queue.
5. Parametric Estimation of Impulse Response Function

We now consider a family $g$ of impulse response functions $g$ given by

$$
g(t) = \begin{cases} 
\beta e^{-\gamma t}, & t > 0 \\
0, & t \leq 0,
\end{cases} \quad (5.1)
$$

with corresponding Fourier transform

$$
\hat{g}(\omega) = \frac{\beta}{\gamma - i\omega}.
$$

$$
= \frac{\beta\gamma + i\beta\omega}{\gamma^2 + \omega^2}. \quad (5.2)
$$

We first consider a model of order 1 (i.e. only linear terms) without noise:

$$
y(t) = \sum_{j=1}^{n} \int_{0}^{\infty} x_j(t - \tau) g_j(\tau) \, d\tau \quad (5.3)
$$

in which each $g_j$ is in $g$; that is, $g_j(t) = \beta_j \exp\{-\gamma_j t\}$. If we write $\hat{g}_j$ in terms of its modulus and argument, we have

$$
\hat{g}_j(\omega) = |\hat{g}_j(\omega)| \exp\{i \varphi_j(\omega) \}.
$$

Thus,

$$
|\hat{g}_j(\omega)|^2 = \frac{\beta_j^2}{\gamma_j^2 + \omega^2} \quad (5.4)
$$

and
\[
\cot \varphi_j(\omega) = \frac{\Re \hat{g}_j(\omega)}{\Re \hat{g}_j(\omega)}
\] (5.5)

If we now could get observations of \( \hat{g}_j \) at frequencies \( \nu_{ij} \), \( i = 1, \ldots, m \), we would have from (5.2)

\[
\frac{\Re \hat{g}_j(\nu_{ij})}{\Im \hat{g}_j(\nu_{ij})} = \gamma_j/\nu_{ij}
\] (5.6)

\[
| \hat{g}_j(\nu_{ij}) |^2 = \frac{| \beta_j |^2}{(\gamma_j^2 + \nu_{ij}^2)}.
\] (5.7)

Thus, method of moments estimators for \( \gamma \) and \( |\beta| \) are:

\[
\hat{\gamma}_j = \frac{1}{m} \sum_{i=1}^{m} \frac{\Re \hat{g}_j(\nu_{ij})}{\Im \hat{g}_j(\nu_{ij})}
\] (5.8)

\[
| \hat{\beta}_j |^2 = \frac{1}{m} \sum_{i=1}^{m} \left[ \hat{\gamma}_j^2 + \nu_{ij}^2 \right] | \hat{g}_j(\nu_{ij}) |.
\] (5.9)

Now let us consider the response of a simple linear system to (scalar) input \( x(t) \) having Fourier sine series

\[
x(t) = \sum_{k=1}^{\infty} b_k \sin k\omega_0 t
\] (5.10)

so that the response \( y(t) \) is
\[ y(t) = \int_0^\infty \sum_{k=1}^\infty b_k \sin k\omega_0(t-\tau) \beta e^{-\gamma \tau} \, d\tau \]

\[ = \sum_{k=1}^\infty \beta b_k \left[ \sin k\omega_0 t \int_0^\infty \cos k\omega_0 \tau e^{-\gamma \tau} \, d\tau 
+ \cos k\omega_0 t \int_0^\infty \sin k\omega_0 \tau e^{-\gamma \tau} \, d\tau \right] \]

\[ = \sum_{k=1}^\infty \frac{\beta b_k}{\gamma^2 + (k\omega_0)^2} \left[ \gamma \sin k\omega_0 t + k\omega_0 \cos k\omega_0 t \right]. \]

Thus, the sine and cosine coefficients of the Fourier series for \( y(t) \) are

\[ \frac{\gamma \beta b_k}{\gamma^2 + (k\omega_0)^2} = \frac{2}{T} \int_0^T y(t) \sin k\omega_0 t \, dt \quad (5.12) \]

and

\[ \frac{k \omega_0 \beta b_k}{\gamma^2 + (k\omega_0)^2} = \frac{2}{T} \int_0^T y(t) \cos k\omega_0 t \, dt. \quad (5.13) \]

Equations (5.12) and (5.13) offer another way of estimating the parameters \( \beta \) and \( \gamma \). Taking the ratios of the two equations, we obtain:

\[ \frac{\hat{\gamma}}{k\omega_0} = \frac{\frac{2}{T} \int_0^T y(t) \sin k\omega_0 t \, dt}{\frac{2}{T} \int_0^T y(t) \cos k\omega_0 t \, dt} \quad (5.14) \]

and squaring equations (5.12) and (5.13) and summing over \( k \) gives
\[ \sum_{k=1}^{\infty} \frac{b_k^2}{\gamma^2 + (\omega_0)^2} = \sum_{k=1}^{\infty} \left[ \frac{2}{T} \int_0^T y(t) \sin k\omega_0 t \, dt \right]^2 + \left[ \frac{2}{T} \int_0^T y(t) \cos k\omega_0 t \, dt \right]^2. \] (5.15)

Now assume that \( x(t) \) is chosen so that

\[ \int_0^T x(t) \, dt = 0 \]

and therefore

\[ \int_0^T y(t) \, dt = 0. \]

Then by Parseval's formula, the right side of equation (5.15) is \( \int_0^T y^2(t) \, dt \).

We can finally estimate \( |\beta| \) using

\[ |\hat{\beta}|^2 = \frac{\int_0^T y^2(t) \, dt}{\sum_{k=1}^{\infty} \frac{b_k^2}{\gamma^2 + (\omega_0)^2}}. \] (5.16)

5.2. **An Example: \( M/M/\infty \) Queue**

We now present a simple example in which the mean number in the system is approximately a linear filter with impulse response function in \( q \).

Consider an \( M/M/\infty \) queue with non-homogeneous arrival and service rates \( \lambda(t) \) and \( \mu(t) \), respectively. The mean number in the system at time \( t \) is

\[ E[N(t)] = \int_0^t \lambda(s) \exp \left\{ - \int_s^t \mu(u) \, du \right\} \, ds \] (5.17)
as shown by the following argument. Conditional on \( n \) arrivals by time \( t \), the
times of arrivals are independent identically distributed random variables
with density

\[
f(s) = \frac{\lambda(s)}{\int_0^t \lambda(u) \, du}.
\tag{5.18}
\]

If an arrival occurs at time \( s \), then it is still in the system at time \( t \)
provided the service time exceeds \( t-s \), and this occurs with probability

\[
\exp\{-\int_s^t \mu(u) \, du\}.
\tag{5.19}
\]

Hence, given \( n \) arrivals in \([0,t]\), the number remaining at time \( t \) is binomial
with parameters \( n \) and

\[
p = \int_0^t f(s) \exp\{-\int_s^t \mu(u) \, du\} \, ds
\tag{5.20}
\]

where \( f(s) \) is given in equation (5.18). Thus, the conditional mean number in
the system is \( np \). Since the mean number of arrivals in \([0,t]\) is

\[
E[ \text{arrivals in } [0,t] ] = \int_0^t \lambda(u) \, du
\tag{5.21}
\]
equation (5.17) follows.

Now let the arrival and service rates be \( \lambda_0 + \lambda(t) \) and \( \mu_0 + \mu(t) \)
respectively, where \( \lambda_0 \) and \( \mu_0 \) are the nominal values with \( \lambda(t) \) and \( \mu(t) \) the
time-varying oscillations about the nominal values. Equation (5.17) now becomes

\[ E[N(t)] = \int_0^t [\lambda_0 + \lambda(s)] \exp\left\{ - \int_s^t [\mu_0 + \mu(u)] \, du \right\} \, ds. \]  

(5.22)

If we now approximate \( \exp\{ - \int \mu(u) \, du \} \) in equation (5.22) by \( 1 - \int \mu(u) \, du \), the corresponding linearized mean is

\[ E[N(t)] = \int_0^t [\lambda_0 + \lambda(s)] \exp\{ - \mu_0(t-s) \} \left[ 1 - \int_s^t \mu(u) \, du \right] \]  

(5.23)

\[ = \lambda_0 \int_0^t \exp\{ - \mu_0(t-s) \} \, ds + \int_0^t \lambda(s) \exp\{ - \mu_0(t-s) \} \, ds \]

\[ - \lambda_0 \int_0^t \exp\{ - \mu_0(t-s) \} \left[ \int_s^t \mu(u) \, du \right] \, ds \]  

(5.24)

+ a higher order term in \( \lambda(s) \) and \( \mu(s) \).

Interchanging the order of integration in the last integral of equation (5.24), we get

\[ \int_0^t \exp\{ - \mu_0(t-s) \} \int_s^t \mu(u) \, du \, ds = \int_0^t \mu(u) \int_0^u \exp\{ - \mu_0(t-s) \} \, ds \, du \]

\[ = \frac{1}{\mu_0} \int_0^t \exp\{ - \mu_0(t-u) \} \mu(u) \, du \]  

(5.25)

\[ - \frac{1}{\mu_0} \exp(-\mu_0 t) \int_0^t \mu(u) \, du. \]

Thus, the mean number in the system is approximately
\[
\frac{\lambda_0}{\mu_0} + \int_0^t \exp\{-\mu_0(t-s)\} \left[ \lambda(s) - \left(\frac{\lambda_0}{\mu_0}\right)\mu(s) \right] \, ds \\
- \left(\frac{\lambda_0}{\mu_0}\right) \exp(-\mu_0 t) \left[ 1 + \int_0^t \mu(u) \, du \right].
\]

5.3. Frequency Selection

In the preceding section as well as in Section 4 it was necessary to obtain estimates of \( g_\alpha(\omega) \) or \( G_\alpha(\omega) \) for the various terms in the model. In this section we discuss one way of obtaining such estimates in a few simulation runs by oscillating parameters as we did for the factor screening experiments.

First, suppose that we have a linear model

\[
y(t) = \sum_{j=1}^n \int_0^\infty x_j(t-\tau) \, g_j(\tau) \, d\tau
\]

as in equation (5.3), and let the input be expressed as a Fourier series

\[
x_j(t) = \sum_{k=1}^\infty b_{jk} \sin k\omega_j t
\]

for \( j=1,\ldots,n \). For practical purposes we are restricted to choosing rational frequencies \( \omega_j \). Thus, for each \( i \neq j \) there are (infinitely many) integers \( k \) and \( l \) such that \( k\omega_i = l\omega_j \). Thus, it is clear that the Fourier series we choose should be finite. Furthermore, the frequencies \( \omega_1,\ldots,\omega_n \) should be chosen to have distinct prime numbers for denominators. To illustrate this, take \( n = 2 \) and let \( \omega_1 = 1/5 \), \( \omega_2 = 1/7 \). Then there will be four multiples of \( 1/5 \) and six of \( 1/7 \) which do not confound, but since \( 5\omega_1 = 7\omega_2 \) there will be confounding at \( 1 \). The estimated spectrum (or Fourier transform) at integer frequencies may be simply ignored in fitting the model. On the other hand, if we simply
omit integer frequencies \((5n\omega_1\) for the first factor, for example) then there will be no confounding at all. This is perhaps a better approach, since all the spectral information will then be concentrated at useful frequencies. Furthermore, the effect of several factors contributing to particular frequencies may tend to overshadow the desired spectral estimates.

Observe that inputs such as in equation (5.28) are somewhat more general than the oscillations of factors for the factor screening experiments in Section 2. However, there is no reason that such inputs could not be used for factor screening experiments. They might be useful to mitigate effects of gain, albeit at the expense of bandwidth (and hence longer runs).

For example, consider a quadratic model. We seek to choose frequencies \(\omega_1, \ldots, \omega_n\) so that the term indicator frequencies give a fine enough grid to accurately approximate \(G_\alpha(\omega)\) with no confounding. If we restrict the terms in the Fourier series to be odd multiples of the respective fundamental frequency, then there will be no confounding between first order and second order terms. This is because all numerators for second order terms (quadratic and interaction) will be even, while the those of linear terms will be odd. Furthermore, since the denominators are prime numbers which are different for each factor, there will be no confounding between linear terms, between quadratic terms for distinct factors, or between interaction terms for distinct pairs of factors. The primary concern now is to avoid confounding between interaction terms and quadratic terms for the same pair of factors. Of course, we also wish to choose frequencies which maximize bandwidth, the minimum distance between adjacent term indicator frequencies. This can be minimized by choosing the numerators for \(\omega_j\)'s to be distinct primes which are also distinct from the denominators. If the denominator for a frequency is \(2m+1\), then only the first \(m\) odd multiples of that frequency
are taken in the Fourier series. Since all numerators and denominators are
distinct prime numbers, there will be no confounding between the interaction
terms and the quadratic terms, and thus no confounding at all. Of course, if
the experimenter believes that no interaction is present, say as the result
of an initial factor screening experiment, then choosing frequencies as in
the linear case above will suffice.

We will illustrate with a model with two factors \( n = 2 \). Table 5.1
shows the term indicator numerators for \( \omega_1 = 3/11 \) and \( \omega_2 = 5/13 \).
Table 5.1 Quadratic Model, Two Factors. Table Gives Term Indicator Numerators. $\omega_1 = 3/11$, $\omega_2 = 5/13$. Denominator = 143

<table>
<thead>
<tr>
<th>Quadratic $\downarrow$</th>
<th>78</th>
<th>234</th>
<th>390</th>
<th>546</th>
<th>702</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>39</td>
<td>117</td>
<td>195</td>
<td>273</td>
<td>351</td>
</tr>
<tr>
<td>Linear $\uparrow$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>172</td>
<td>250</td>
<td>328</td>
<td>406</td>
</tr>
<tr>
<td>110</td>
<td>16</td>
<td>62</td>
<td>140</td>
<td>218</td>
<td>296</td>
</tr>
<tr>
<td></td>
<td>204</td>
<td>282</td>
<td>360</td>
<td>438</td>
<td>516</td>
</tr>
<tr>
<td>330</td>
<td>126</td>
<td>48</td>
<td>30</td>
<td>108</td>
<td>186</td>
</tr>
<tr>
<td></td>
<td>314</td>
<td>392</td>
<td>470</td>
<td>548</td>
<td>626</td>
</tr>
<tr>
<td>550</td>
<td>236</td>
<td>158</td>
<td>80</td>
<td>2</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>424</td>
<td>502</td>
<td>580</td>
<td>658</td>
<td>736</td>
</tr>
<tr>
<td>770</td>
<td>346</td>
<td>268</td>
<td>190</td>
<td>112</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>534</td>
<td>612</td>
<td>690</td>
<td>768</td>
<td>846</td>
</tr>
<tr>
<td>990</td>
<td>456</td>
<td>378</td>
<td>300</td>
<td>222</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>644</td>
<td>722</td>
<td>800</td>
<td>878</td>
<td>956</td>
</tr>
<tr>
<td>1210</td>
<td>566</td>
<td>488</td>
<td>410</td>
<td>332</td>
<td>254</td>
</tr>
</tbody>
</table>
References


Appendix I: Generalized Functions and Processes

There are two places where the concept of generalized functions has been used. One is the use of the Dirac delta function in the Fourier transform pair

\[
\cos 2\pi \omega_0 t \leftrightarrow \frac{\delta(\omega - \omega_0) + \delta(\omega + \omega_0)}{2}
\]  

(I.1)

used in the spectral analysis in section 2. To demonstrate this we need a few results from Fourier analysis of tempered distributions. The other use of generalized functions is in the notion of "white noise" processes. In continuous time they cannot be defined as an ordinary stochastic process. However, a generalized random process, a random element with values in the space of generalized functions, may have the desired "flat" spectrum.

The terms "distribution" and "generalized function" are both used in the literature. The former tends to be avoided by probabilists because of the connotation of distribution of a random variable. However, Fourier analysis on generalized functions is traditionally done in the Schwartz space of "tempered distributions". Therefore, our compromise will be to refer to tempered distributions while doing (deterministic) Fourier analysis and to refer to generalized processes as random elements in the space of distributions.

I.1. Heuristic Treatment of the Dirac Delta Function

The Dirac delta function \(\delta\) has the property that for all functions \(f\) continuous at the origin

\[
\int_{-\infty}^{\infty} f(t) \delta(t) \, dt = f(0).
\]  

(I.2)
Such an object does not exist as an ordinary function; nevertheless, it may
be used formally according to equation (I.2) as if it were and the answer
will (almost always) be correct. So long as δ(t) only exists inside
integrals of the form of the left side of equation (I.2), in fact, there will
be no difficulties.

Certain operations may be performed on δ as if it were an ordinary
function. For example, a shift in its argument is handled using (formal)
change of variables within the integral:

\[ \int_{-\infty}^{\infty} f(t) \, \delta(t - a) \, dt = \int_{-\infty}^{\infty} f(t + a) \, \delta(t) \, dt = f(a), \number (I.3) \]

providing \( f \) is continuous at \( a \). We can perform the convolution of δ with an
ordinary function \( f \):

\[ (\delta \ast f)(t) = \int_{-\infty}^{\infty} \delta(\tau) \, f(t - \tau) \, d\tau = f(t). \number (I.4) \]

The Fourier transform \( \hat{\delta} \) of δ is the function identically 1, which we will
denote by 1:

\[ \hat{\delta}(\omega) = \int_{-\infty}^{\infty} e^{2\pi i \omega t} \, \delta(t) \, dt = e^{2\pi i \omega(0)} = 1. \number (I.5) \]

The usual rules for Fourier transforms hold. For example, the transform of
the convolution is the product of the transformations:

\[ (\delta \ast f)^{(\omega)} = \int_{-\infty}^{\infty} e^{2\pi i \omega t} \, dt \int_{-\infty}^{\infty} \delta(\tau) \, f(t - \tau) \, d\tau \]
\[
\hat{\delta}(\omega) \cdot \hat{f}(\omega) = \int_{-\infty}^{\infty} e^{2\pi i \omega t} \hat{f}(\omega) \, dt
\]

Also, a shift in the argument leads to multiplying the transform by an exponential:

\[
[\delta(t-\tau)](\omega) = \int_{-\infty}^{\infty} e^{2\pi i \omega(t-\tau)} \delta(t) \, dt
\]
\[
= \int_{-\infty}^{\infty} e^{2\pi i \omega(t+\tau)} \delta(t) \, dt
\]
\[
= e^{2\pi i \omega \tau} \hat{\delta}(\omega) = e^{2\pi i \omega \tau}.
\]

Now recall the inversion formula for the Fourier transform of an ordinary function is:

\[
f(t) = \int_{-\infty}^{\infty} e^{-2\pi i \omega t} \hat{f}(\omega) \, d\omega,
\]

providing, for example, that both \( f \) and \( \hat{f} \) are absolutely integrable (Stein and Weiss [1971]). Taking for granted that the inversion formula holds for expressions involving \( \delta \), we have:

\[
\delta(t) = \hat{1}(-\omega) = \hat{1}(\omega).
\]

Thus, using equation (I.7) we have
\[ [e^{2\pi i \eta t}]^\sim(\omega) = \hat{1}(\omega - \eta) = \delta(\omega - \eta). \] (I.9)

Equation (I.9) gives the desired result for the spectrum of oscillated factors. We have:

\[
(\cos 2\pi \eta t)^\sim(\omega) = \left( \frac{1}{2} \left[ e^{2\pi i \eta t} + e^{-2\pi i \eta t} \right] \right)^\sim(\omega).
\]

\[ = \frac{1}{2} \left[ \delta(\omega - \eta) + \delta(\omega + \eta) \right]. \] (I.10)

Thus, the power spectrum of a stationary process whose covariance function is sinusoidal with frequency \(\eta\) consists of "spikes" at frequencies \(\pm \eta\). Because of symmetry, it is usual to only graph the power spectrum only for positive frequencies.

A rigorous development of the Dirac delta function requires the concept of generalized functions (distributions), which are presented in the following sections. The material which follows is not necessary to perform frequency domain experiments, but is simply a somewhat rigorous exposition of the concepts needed to make generalized functions and processes precise. The following leans rather heavily on Rudin (1976) and Stein and Weiss (1971), to which the reader is referred for further details. A good heuristic approach is contained in Papoulis (1963).

I.2. Test Functions

To put distributions on rigorous footing Laurent Schwartz defined them as linear functionals on certain spaces of "test functions". These are spaces consisting of sufficiently well-behaved functions, on which all desired operations may be performed (e.g. differentiation, taking Fourier
transforms, etc.). Accordingly, the space $\mathcal{S}$ consists of all $C^\infty$ functions $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$\sup_x |x^\alpha D^\beta \varphi(x)| < \infty$$  

(I.11)

for all multi-indices $\alpha, \beta$, where $D = (\partial/\partial x_1, \ldots, \partial/\partial x_n)$. One function in $\mathcal{S}$ is $\exp(-|x|^2)$. However, the function $\exp(-|x|)$ is not differentiable at the origin, and hence not in $\mathcal{S}$.

To define continuous linear functional on $\mathcal{S}$ we need to topologize $\mathcal{S}$. It turns out that $\mathcal{S}$ can be made into a complete separable metric space. Define the family of metrics $\{\rho_{\alpha\beta}(\varphi, \psi)\}$ by

$$\rho_{\alpha\beta}(\varphi, \psi) = \sup_x |x^\alpha D^\beta (\varphi - \psi)|.$$  

(I.12)

Let $\{\rho_n\}$ be an ordering of this countable family of metrics and put

$$d(\varphi, \psi) = \sum_{n=1}^{\infty} 2^{-n} \min \{ \rho_n(\varphi, \psi), 1 \}$$  

(I.13)

In equation (I.13), $d$ is a metric on $\mathcal{S}$. Furthermore, a sequence $\{\varphi_k\}$ converges to $\varphi$ in the metric $d$ if and only if $\varphi_k \rightarrow \varphi$ with respect to each metric $\rho_n$ (i.e. $\rho_n(\varphi_k, \varphi) \rightarrow 0$ as $k \rightarrow \infty$, $\forall n$).

I.3. Tempered Distributions

A linear functional on $\mathcal{S}$ is a mapping $\psi: \mathcal{S} \rightarrow \mathbb{R}$ such that

$$\psi(\alpha_1 \varphi_1 + \alpha_2 \varphi_2) = \alpha_1 \psi(\varphi_1) + \alpha_2 \psi(\varphi_2)$$  

(I.14)

for all $\alpha_1, \alpha_2 \in \mathbb{R}$ and $\varphi_1, \varphi_2 \in \mathcal{S}$. $\psi$ is a continuous linear functional if it
is a linear functional that is continuous in the topology induced by the metric $d$ above. Continuity at the origin (zero function) is necessary and sufficient for a linear functional to be continuous. The space $\mathcal{S}'$ of continuous linear functionals on $\mathcal{S}$ is called the space of tempered distributions. As the following examples illustrate, the word "tempered" refers to a restriction on the growth at infinity (see Rudin, 1976). $\mathcal{S}'$ will be endowed with the weak-∗ topology as follows.

To each $\varphi \in \mathcal{S}$ there corresponds a linear functional $\Lambda_{\varphi}$ on $\mathcal{S}'$ defined by

$$\Lambda_{\varphi} \psi = \varphi(\psi), \; \psi \in \mathcal{S}. \quad (I.15)$$

The weak-∗ topology on $\mathcal{S}'$ is the smallest topology on $\mathcal{S}'$ such that all the functionals $\{\Lambda_{\varphi}, \varphi \in \mathcal{S}\}$ are continuous. Important examples of tempered distributions are the following.

i) Bounded Functions

Let $\sup |f(x)| < \infty$ and define $\psi_f: \mathcal{S} \rightarrow \mathbb{R}$ by $\psi_f(\varphi) = \int f(x)dx$. Then $\sup (1+x^2)^N |\varphi_n(x)| \rightarrow 0$ if $\varphi_n \rightarrow 0$ by the topology defined on $\mathcal{S}$. Therefore,

$$| \psi_f(\varphi_n) | \leq \|f\|_{\infty} \sup (1+x^2)^N |\varphi_n(x)| \int \frac{dx}{(1+x^2)^N} \rightarrow 0 \; (n \rightarrow \infty) \quad (I.16)$$

Thus, $\psi_f \in \mathcal{S}'$. Similarly, each $f \in L^p \; (p \geq 1)$ may be associated with a tempered distribution.

ii) Finite Borel Measures

Let $\mu$ be a finite Borel measure on $\mathbb{R}^n$ and define $\psi_\mu: \mathcal{S} \rightarrow \mathbb{R}$ by

$$\psi_\mu(\varphi) = \int \varphi(x) \mu(dx). \; \psi_\mu \text{ is clearly linear, and continuity follows from the estimate}$$
\[ |\varphi_\mu(\varphi)| \leq \|\varphi\|_{\infty} \mu(\mathbb{R}^n). \quad (I.17) \]

iii) Slowly Increasing Functions

A real-valued function \( f \) on \( \mathbb{R}^n \) is said to be slowly increasing if \[ \sup |f(x)|/(1+|x|^2)^k \to \infty \text{ for some } k. \] For such a function \( \varphi_f \) as defined in i) above is a tempered distribution.

iv) Dirac Delta Function

Let \( \varphi: \mathcal{D} \to \mathbb{R}^n \) be defined by \( \varphi(\varphi) = \varphi(0) \); then \( \varphi \in \mathcal{D}' \). This is suggestively written as \[ \int \delta(x) \varphi(x) \, dx = \varphi(0), \] where \( \delta \) is the "function" associated with \( \varphi \), called the Dirac delta function. Clearly no ordinary function can have such a property. However, let \( \mu \) be the measure putting mass 1 at 0. Then \[ \int \varphi(x) \mu(dx) = \varphi(0) \text{ for all } \varphi \in \mathcal{D}. \] Thus, \( \delta \) may be associated with \( \mu \) as in ii).

In general tempered distributions may not be multiplied together. However, they may be multiplied with elements of \( \mathcal{D} \); if \( \varphi \in \mathcal{D} \) and \( \varphi \in \mathcal{D}' \), then \( \varphi \varphi \) is defined by

\[ \varphi \varphi(\nu) = \varphi(\varphi \nu) \]

I.4. Fourier Transform of Tempered Distributions

For a function \( \varphi \in \mathcal{D} \) the Fourier transform may be defined as

\[ \hat{\varphi}(\omega) = \int_{\mathbb{R}^n} \varphi(x) e^{2\pi i (\omega,x)} \, dx. \quad (I.18) \]

This is well-defined since \( \mathcal{D} \subset L^1 \) (i.e. \( \int |\varphi(x)| \, dx < \infty \) for \( \varphi \in \mathcal{D} \)).
Furthermore, the inversion formula holds:

\[ \varphi(x) = \int_{\mathbb{R}^n} \hat{\varphi}(\omega) e^{-2\pi i (\omega, x)} d\omega. \]  \hspace{1cm} (I.19)

It turns out that \( \hat{\varphi} \in \mathcal{S} \). We first note that

\begin{align*}
\text{a)} & \quad D^\beta \hat{\varphi}(\omega) = [(2\pi i x)^\beta \varphi(x)]^- \omega \\
\text{b)} & \quad [x^\alpha \varphi(x)]^- \omega = (2\pi i \omega)^\alpha \hat{\varphi}(\omega). \hspace{1cm} (I.20)
\end{align*}

Thus,

\[ \sup_{\omega^\alpha} | \omega^\alpha D^\beta \hat{\varphi}(\omega) | = \sup_{\omega^\alpha} \left| (2\pi i)^\beta \varphi(x) \right|^\omega (\omega) \mid \\
= (2\pi)^{|\beta|} |\alpha| \sup_{\omega^\alpha} \left| D^\alpha x^\beta \varphi(x) \right| \mid \\
\leq (2\pi)^{|\beta|} |\alpha| \| D^\alpha x^\beta \varphi(x) \|_1 < \omega. \]

Therefore, \( \hat{\varphi} \in \mathcal{S} \). We can now define the Fourier transform of a tempered distribution \( \varphi \in \mathcal{S}' \) by

\[ \hat{\psi}(\varphi) = \psi(\hat{\varphi}), \varphi \in \mathcal{S}. \]  \hspace{1cm} (I.21)

That is, \( \hat{\psi} \) is the linear functional on \( \mathcal{S} \) which assigns the value \( \psi(\hat{\varphi}) \) to each function \( \varphi \in \mathcal{S} \). Observe that the fact that the Fourier transform maps \( \mathcal{S} \) into \( \mathcal{S} \) is crucial for this definition; otherwise the right hand side of equation (I.21) is not defined.

The motivation for this definition comes from distributions associated with functions, as in i) of section 2. Let \( f \in L^1 \); the tempered distribution associated with \( \hat{f} \) is given by
\[ \psi_f(\varphi) = \int \hat{f}(x) \varphi(x) \, dx \]
\[ = \int \varphi(x) \, dx \int e^{2\pi i(x,t)} f(t) \, dt \]
\[ = \int f(t) \, dt \int e^{2\pi i(x,t)} \varphi(x) \, dx \]
\[ = \int f(t) \hat{\varphi}(t) \, dt \]
\[ = \psi_f(\hat{\varphi}). \]

We will now obtain the transform pair used in the spectral analysis of the DRSW. First, we shall find the Fourier transform of the Dirac delta function. We have

\[ \hat{\delta}(\varphi) = \delta(\hat{\varphi}) = \hat{\varphi}(0) = \int \varphi(x) \, dx. \]  

(1.22)

Let 1 be the element of \( \mathcal{L}' \) defined by \( 1(\varphi) = \int \varphi(x) \, dx \). Then (1.13) shows that \( \hat{\delta} = 1 \). Conversely, the Fourier transform of 1 is

\[ \hat{1}(\varphi) = 1(\hat{\varphi}) = \int \hat{\varphi}(\omega) \, d\omega = \varphi(0) = \delta(\varphi). \]  

(1.23)

The third equality in (1.14) is by the inversion formula. To obtain the final result, we need to introduce the shift function \( \tau_x: \mathcal{L} \rightarrow \mathcal{L} \) by \( \tau_x \varphi(y) = \varphi(y-x) \). This is then used to define a shift on \( \mathcal{L}' \) by

\[ (\tau_x \varphi)(\varphi) = \psi(\tau_{-x} \varphi). \]  

(1.24)

Observe that
a) \( \tau_{\lambda} \hat{\varphi}(\omega) = \left[ e^{-2\pi i (\lambda, x)} \varphi(x) \right]^-(\omega) \)

\( (I.25) \)

b) \( e^{2\pi i (\omega, x)} \hat{\varphi}(\omega) = [\tau_{x} \varphi(x)]^-(\omega) \quad \varphi \in \mathcal{S} \).

We can derive similar identities for tempered distributions as follows.

\[
\tau_{\lambda} \hat{\varphi}(\varphi) = \hat{\varphi}(\tau_{-\lambda} \varphi) \\
= \varphi([\tau_{-\lambda} \varphi]^+) \\
= \varphi(e^{-2\pi i (\lambda, \cdot)} \hat{\varphi}) \\
= e^{-2\pi i (\lambda, \cdot)} \varphi(\hat{\varphi}) \\
= [e^{-2\pi i (\lambda, \cdot)} \varphi]^-(\varphi).
\]

Also,

\[
e^{2\pi i (\lambda, \cdot)} \hat{\varphi}(\varphi) = \hat{\varphi}(e^{2\pi i (\lambda, \cdot)} \varphi) \\
= \varphi([e^{2\pi i (\lambda, \cdot)} \varphi]^+) \\
= \varphi(\tau_{-\lambda} \hat{\varphi}) \\
= \tau_{\lambda} \varphi(\hat{\varphi}) \\
= [\tau_{\lambda} \varphi]^-(\varphi)_{\lambda}
\]

To summarize, for \( \varphi \in \mathcal{S}' \), we have

\[
a) \quad \tau_{\lambda} \hat{\varphi} = [e^{-2\pi i (\lambda, \cdot)} \varphi]^+ \\
b) \quad e^{2\pi i (\lambda, \cdot)} \hat{\varphi} = [\tau_{\lambda} \varphi]^-
\]

\( (I.26) \)

We now take \( \varphi = 1 \) in equation \( (I.26a) \) and obtain
\[ [e^{-2\pi i (\lambda, \cdot)}]^\text{t} = \tau_{\lambda}^\delta \]  

(I.27)

Therefore,

\[ [\cos 2\pi (\lambda, \cdot)]^\text{t} = [\frac{1}{2} (e^{2\pi i (\lambda, \cdot)} + e^{-2\pi i (\lambda, \cdot)})]^\text{t} = \frac{1}{2} [\tau_{\lambda}^\delta + \tau_{-\lambda}^\delta]. \]  

(I.28)

This is the desired result, as can be seen by writing \( \delta \) in its suggestive "function" notation \( \delta(\omega) \). Then, equation (I.28) becomes

\[ [\cos(2\pi(\lambda, \cdot))]^\text{t}(\omega) = \frac{1}{2} [\delta(\omega - \lambda) + \delta(\omega + \lambda)] \]  

(I.29)

I.5. Differentiation of Tempered Distributions

Let \( D = (\partial/\partial x_1, \ldots, \partial/\partial x_n) \) and let \( \alpha \) be a multi-index. For a tempered distribution \( \psi \in \mathcal{S}' \) the derivative \( D^\alpha \psi \) is defined to be

\[ [D^\alpha \psi](\varphi) = (-1)^{|\alpha|} \psi(D^\alpha \varphi). \]  

(I.30)

This is again motivated by those tempered distributions associated with functions. For example, using the prime notation for derivatives on the real line, the derivative of the \( \delta \) function is given by

\[ [\delta'](\varphi) = -\delta(\varphi) = -\varphi'(0). \]

Since the right side of equation (I.30) is defined for all \( \varphi \in \mathcal{S} \), we see that all tempered distributions have derivatives of all orders.

I.6. Generalized Processes
Let $B$ be the Borel $\sigma$-algebra on $\mathcal{S}'$ generated by the weak-*$\sigma$ topology. If $(\Omega, \mathcal{F}, P)$ is a probability space, then a generalized stochastic process is a mapping $\xi: \Omega \rightarrow \mathcal{S}'$ that is $\mathcal{F}/B$-measurable.

If $\{X(t)\}$ is an ordinary stochastic processes then it may be identified with a generalized stochastic processes in the same manner as functions are identified with distributions; to be in $\mathcal{S}'$ the sample paths of $\{X(t)\}$ must meet growth requirements. For example, if there is some $k$ such that almost surely we have

$$\sup_{t} \frac{X(t)}{[1 + t^2]^k} < \infty \quad (1.31)$$

Then $X(t)$ may be identified with $\Psi_X \in \mathcal{S}'$ defined by $\Psi_X(\phi) = \int X(t)\phi(t)dt$ for all $\phi \in \mathcal{S}$. The version of $X(t)$ is assumed to be measurable so that the preceding integral is defined.

Associated with each generalized process $\psi$ are the mean $m_\psi$, a linear functional on $\mathcal{S}$, and covariance $R_\psi$, which is a bilinear form on $\mathcal{S} \times \mathcal{S}$. These are defined by

$$m_\psi(\phi) = \mathbb{E}[\psi(\phi)] \quad (\phi \in \mathcal{S}) \quad (1.32)$$

$$R_\psi(\phi, \theta) = \mathbb{E}[\psi(\phi) - m_\psi(\phi)][\psi(\theta) - m_\psi(\theta)] \quad (\phi, \theta \in \mathcal{S}) \quad (1.33)$$

provided the indicated expectations exist. Observe that $\psi(\phi)$ is a real-valued random variable.

A generalized process $\psi$ is Gaussian if for all $\phi_1, \ldots, \phi_n \in \mathcal{S}$, $(\psi(\phi_1), \ldots, \psi(\phi_n))$ is a Gaussian vector. As with ordinary Gaussian processes, generalized Gaussian processes are characterized by their mean and covariance. An ordinary Gaussian processes with regular enough sample paths
may be identified with a generalized Gaussian processes. In such a case, let \( \mu(t) \) and \( C(s,t) \) be, respectively, the mean and covariance functions of the ordinary Gaussian process. Then the mean and covariance functionals with the associated generalized Gaussian process are:

\[
m(\varphi) = E\Psi_X(\varphi) = E \int X(t) \varphi(t) \, dt
\]
\[= \int \mu(t) \varphi(t) \, dt \tag{I.34}
\]

\[
R(\varphi, \Theta) = \int \int C(s,t) \varphi(s) \Theta(t) \, ds \, dt. \tag{I.35}
\]

The measurability of \( X(t) \) is used to invoke Fubini's theorem in equations (I.34-I.35).

We now observe that since all tempered distributions have derivatives of all orders, so do all generalized functions. Denote the derivative of a generalized process \( \Psi \) by \( \dot{\Psi} \). The mean and covariance of \( \dot{\Psi} \) are

\[
m_{\Psi}'(\varphi) = E\dot{\Psi}(\varphi) = -E\Psi(\varphi') = -m_{\Psi}(\varphi') = m_{\Psi}'(\varphi) \tag{I.36}
\]

\[
R_{\Psi}'(\varphi, \Theta) = E[\Psi(\varphi') - m_{\Psi}(\varphi')] [\Theta(\varphi') - m_{\Psi}(\Theta')]
\]
\[= R_{\Psi}(\varphi', \Theta')
\]
\[= D^{(1,1)} R_{\Psi}(\varphi, \Theta). \tag{I.37}
\]

We will now show that Gaussian white noise may be regarded as the derivative of Brownian motion (in the generalized sense). We seek a generalized Gaussian process with mean zero and power spectrum \( f(\omega) = \sigma^2 \). Equivalently, the process should be "delta-correlated"; that is, its correlation function should be \( R(\tau) = \sigma^2 \delta(\tau) \). We first identify Brownian
motion \{B(t)\} with its corresponding generalized process. The law of the iterated logarithm implies that (almost surely)

\[
\sup_t \frac{|B(t)|}{1 + t^2} < \infty.
\]

(I.38)

Therefore, we may consider the sample paths \(t \mapsto B(t)\) to be in \(\mathcal{S}'\) and may thus regard \(B(t)\) as a generalized process, which we will also denote by \(B\). The mean and covariance functionals are

\[
m_B(\varphi) = \int_0^\infty EB(t) \varphi(t) \, dt = 0
\]

(I.39)

\[
R_B(\varphi, \Theta) = \int_0^\infty \int_0^\infty E[B(s)B(t)] \varphi(s) \Theta(t) \, ds \, dt
\]

\[
= \int_0^\infty \int_0^\infty \sigma^2 \min\{s, t\} \varphi(s) \Theta(t) \, ds \, dt
\]

The mean and covariance functionals of the derivative \(\dot{B}\) are therefore

\[
m_B'(\varphi) = m_B'(\varphi) = 0
\]

\[
R_B'(\varphi, \Theta) = R_B'(\varphi', \Theta')
\]

\[
= \int_0^\infty \int_0^\infty \sigma^2 \min\{s, t\} \varphi'(s) \Theta'(t) \, ds \, dt
\]

\[
= \sigma^2 \int_0^\infty \Theta'(t) \, dt \left[ \int_0^t s \varphi'(s) \, ds + \int_t^\infty \varphi'(t) \, ds \right]
\]

\[
= \sigma^2 \int_0^\infty \Theta'(t) \, dt \left[ t \varphi(t) - \int_0^t \varphi(s) \, ds - t \varphi(t) \right]
\]
\[
\begin{align*}
&= -\sigma^2 \int_0^\infty \vartheta'(t) \, dt \int_0^t \varphi(s) \, ds \\
&= -\sigma^2 \int_0^\infty \varphi(s) \, ds \int_s^\infty \vartheta'(t) \, dt \\
&= \sigma^2 \int_0^\infty \varphi(s) \, \Theta(t) \, ds \, dt \\
&= \int_0^\infty \int_0^\infty \sigma^2 \delta(t-s) \varphi(s) \, \Theta(t) \, ds \, dt
\end{align*}
\]

We write this heuristically as

\begin{align*}
E\hat{\vartheta}(t) &= 0 \\
E[\hat{\vartheta}(s) \hat{\vartheta}(t)] &= \sigma^2 \delta(s - t).
\end{align*}

Finally, for \( \varphi_1, \ldots, \varphi_n \in \mathcal{F} \), we have

\[
(B(\varphi_1), \ldots, B(\varphi_n)) = -(B(\varphi_1), \ldots, B(\varphi_n))
\]

so that \( \hat{\vartheta} \) is a Gaussian generalized process.

Another delta-correlated process may be obtained by taking the derivative of a compound Poisson process. Let \( \{N(t)\} \) be a Poisson process with rate \( \lambda \) and \( \{Y_k\} \) by iid random variables with zero mean and variance \( \sigma^2 \).

The compound Poisson process \( \{X(t)\} \) defined by

\[
X(t) = \sum_{k=1}^{N(t)} Y_k
\]

has zero mean and covariance
\[ R_X(t) = \sigma^2 \lambda \min\{s,t\}. \quad (I.43) \]

The generalized process associated with \( \{X(t)\} \) has mean functional zero and covariance functional

\[ R_X(\varphi, \theta) = \sigma^2 \lambda \int_0^\infty \int_0^\infty \delta(s-t) \varphi(s) \theta(t) \, ds \, dt. \quad (I.44) \]

If \( \{T_k\} \) are the times of the jumps of \( \{N(t)\} \), then \( \{\dot{X}(t)\} \) may be represented

\[ X(t) = \sum_{T_k \leq t} Y_T \delta(t - T_k). \quad (I.45) \]