DEPARTMENT OF STATISTICS

The University of Wisconsin
Madison, Wisconsin

TECHNICAL REPORT NO. 314

October 1972

D-OPTIMAL DESIGNS FOR DYNAMIC MODELS

Part I. Theory

by

Bernard Viort

Typist: Jacquelyn R. Forer

This research was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-72-2363 and by the Wisconsin Alumni Research Foundation.

D-OPTIMAL DESIGNS FOR DYNAMIC MODELS

The problem considered here is directly related to the development of realistic dynamic models made possible by the methods of Box and Jenkins [3]. This report tries to provide some partial answer to one of the questions raised in their book (see [3], pp 416-420):

(P) Given a dynamic system:

$$y_t = \frac{\omega(B)}{\delta(B)} x_{t-b} + e_t \qquad t = \ldots,-1,0,1,\ldots \tag{1}$$

Suppose that one can choose the input $\{x_t\}$..
What is the best[1] choice for $\{x_t\}$?

Except for the completely solved simple example used by Box and Jenkins to illustrate their point and for the work of Minnich [17], there seems to exist no statistical literature directly related to this problem. Some remarks on the subject may be found in the engineering literature (see Minnich's bibliography for references, as well as Dhrymes, Klein and Steiglitz [6]) but none of them can be considered as a starting point for a satisfying answer.

The approach presented here is in some sense the result of scattered observations in Courrège and Philoche [5], Durbin [7], Fishman [9], Grenander and Rosenblatt [10], Hannan [11], Kruskal [15],

---

[1] What is meant by best will be discussed below.

Parzen [18], Wahba [22] and Watson [23], observations which led to Viort [21] and will hopefully be discussed elsewhere.

The result of [21], namely the dispersion matrix of the maximum likelihood estimators of the ω- and δ-parameters in (1), is used to relate the problem under consideration to a special case of the "classical" theory of experimental designs as developed by, among others Kiefer (and) Wolfowitz, Karlin and Studden ... Complete references on this subject can be found in St. John [20] and will therefore not be discussed here. Many important results and methods are presented and updated in Fedorov [8]; this book can be considered as the basic reference on the subject.

The question of "optimality" is, and will remain, an open question--if only because it is related to too many different aspects of the problem under consideration, from the assumptions underlying the model to the cost of the experimentation. Optimality is to be taken here in its narrow meaning of optimal for a given problem, a given situation, under well specified assumptions.

Corresponding to these remarks, the situation in the problem (P) is complex:

- Dealing with time series models and using the methodology of Box and Jenkins, one thinks immediately of three possible goals of optimality: optimal for identification, for estimation (all or part of the parameters) and/or diagnostic checking.
- If one is interested in the best estimation of some parameters, there are many different estimates and many notions of optimality each with advantages and disadvantages, and not equivalent as soon as one is dealing with multi-parameter problems.

- The model (1) is non-linear.

I. Problem, Model and Assumptions

This report is limited to the following situation:

- One is only interested in the optimal estimation of the ω- and δ-parameters.
- The method of estimation to be used is maximum likelihood estimation.
- The notion of optimality to be used is D-optimality, i.e. a design is D-optimal if the determinant of the dispersion matrix of the estimators is minimal. This criterion is related to the HPD (highest posterior density) criterion which consists in minimizing the area of the Bayesian HPD region for the parameters. As pointed out in Box and Tiao [4], one of its disadvantages is that it reflects only one aspect of the covariance matrix: the product of the eigenvalues.
- Since the problem is non linear, only D-optimal designs at a given point (ω_0,δ_0)[2] will be considered. This has the advantage to make the results presented here useful in two ways
  1) To develop an iterative method in order to find the optimal input for an unknown model (Box and Hunter [2]),
  2) For a given model, to find how good (or robust) is a given input.

---

[2] For notational convenience, the subscript 0 will be omitted in the following way: (ω,δ) is to be considered as known.

For a practical use of the results to be derived here, the above points imply that the model (1) has been completely identified, estimated and checked, using an ad-hoc procedure. Note that this implies a choice for the input $\{x_t\}$, making the results of this report also useful for this preliminary stage. Since the error process will have some importance, one could in addition consider an input such that $\{x_t=0\}$ and the methods of the first part of Box and Jenkins [3] to identify it separately.

Now in the model (1):

$$\begin{cases} \omega(B) = \omega_0 + \omega_1 B + \dots + \omega_p B^p \\ \delta(B) = 1 - \delta_1 B - \dots - \delta_q B^q \end{cases} \tag{2}$$

where $\omega = (\omega_0,\dots,\omega_p)$, $\delta = (\delta_1,\dots,\delta_q)$, p,q are known.

Let

$$m = p + q + 1 \tag{3}$$

be the total number of parameters to be estimated.

The other assumptions are those of [21]:

A1. All the roots of $\omega(z) = 0$, $\delta(z) = 0$ are outside the unit circle $|z| = 1$ in the complex plane C;

A2. There exists no common root $z_0 \varepsilon C$: $\omega(z_0) = \delta(z_0) = 0$;

A4. $\{e_t\}$ is a stationary stochastic process with mean 0 and known spectral density $\sigma_e^2 f_e(\theta)$ bounded away from zero.

The solution $\{x_t\}$ will be looked for in the class of stationary stochastic processes with mean 0 and spectral density[3] $\sigma_x^2 f_x(\theta)$. $\{x_t\}$ will be independent of $\{e_t\}$. Since $\{x_t\}$ and $\{e_t\}$ are real, $f_x(2\pi-\theta) = f_x(\theta)$, $f_e(2\pi-\theta) = f_e(\theta)$, but the results derived here could be applied to non-real processes. Finally note that the forms of the spectral densities imply that f_x and f_e are normalized.

At the end of this introduction it is worth noting that, despite the many limitations and assumptions, not only are none of them completely unrealistic but most of the actual situations where a dynamic model can be used fit into this frame.

---

[3] $f_x(\theta)$ is to be considered as a generalized function. To avoid mathematical difficulties, f_x will be considered as the sum of two functions:

$f_{x;1}$ which is a.e [dx] continuous,

$f_{x;2}$ which is a finite linear combination of δ functions: $\sum_{\rho=1}^{r} a_\rho \delta_{\theta_\rho}$ where $a_\rho > 0$.

By definition, and with some abuse of notation, for g: $[0,2\pi] \to R$,

$$\int_0^{2\pi} g(\theta) f_{x;1}(\theta) d\theta + \sum_{\rho=1}^{r} a_\rho g(\theta_\rho) \text{ will be represented by}$$

$$\int_0^{2\pi} g(\theta) f_x(\theta) d\theta$$

[A rigourous treatment would necessitate the introduction of the spectral distribution function, which is not exactly relevant to the design problem.]

II. The Relationship with the Standard Experimental Design Problem

With some variation in the notations, the model

$$y_\theta = \alpha_1\phi_1(\theta) + \ldots + \alpha_m\phi_m(\theta) + e_\theta \tag{4}$$

where θ is a real "control" variable $(\theta\varepsilon[0,2\pi])$

y_θ is the real observation at θ

$\alpha = (\alpha_1,\ldots,\alpha_m)$ is a real vector of unknown parameters

$\phi_j(\theta)$, $j=1,\ldots,m$ are known real independent functions

e_θ are real independent (for different θ) random variables

with mean 0 and variance $\frac{\sigma_e^2}{\sigma_x^2} f_e(\theta) > 0$

is the starting point of the now classical literature on experimental designs.

If one takes n measurements (corresponding to $\theta_1,\ldots,\theta_n$ not necessarily all different) the information[4] matrix of the parameter α is

$$I_\alpha^{(n)} = \left\{ \sum_{\nu=1}^{n} \phi_j(\theta_\nu)\phi_k(\theta_\nu) \frac{\sigma_x^2}{\sigma_e^2 f_e(\theta_\nu)} \; ; \quad j,k = 1,\ldots,m \right\} \tag{5}$$

and, if there are N different θ's, with n_μ measurements at θ_μ,

$$\sum_{\mu=1}^{N} n_\mu = n \tag{6}$$

---

[4] From now on the inverse of the dispersion matrix will be used, and called the information matrix. The criterion of D-optimality is then to maximize the determinant of the information matrix.

let

$$p_\mu = \frac{n_\mu}{n} \tag{7}$$

then

$$I_\alpha^{(n)} = \left\{ n \sum_{\mu=1}^{N} \phi_j(\theta_\mu)\phi_k(\theta_\mu) \frac{\sigma_x^2 \; p_\mu}{\sigma_e^2 \; f_e(\theta_\mu)} \; ; \quad j,k = 1,\ldots,m \right\} \tag{8}$$

The solution consisting in increasing n, the number of measurement, clearly presents no interest and what one is interested in is maximizing the information "per observation", i.e. the determinant of

$$I_\alpha = \left\{ \sum_{\mu=1}^{N} \phi_j(\theta_\mu)\phi_k(\theta_\mu) \frac{\sigma_x^2 \; p_\mu}{\sigma_e^2 \; f_e(\theta_\mu)} \; ; \quad j,k = 1,\ldots,m \right\} \tag{9}$$

The form of I_α suggests the identification of a design with a (discrete) probability measure on $[0,2\pi]$. It is very convenient, and justified, to drop the requirement of the measure being discrete and identify a design with a probability measure (to be represented by $f_x(\theta)$) on $[0,2\pi]$. Under some mild regularity conditions on the ϕ_j's[5] and with the extended definition of the integral, I_α is then

$$I_\alpha = \left\{ \int_0^{2\pi} \phi_j(\theta)\phi_k(\theta) \frac{\sigma_x^2 \; f_x(\theta)}{\sigma_e^2 \; f_e(\theta)} \, d\theta \; ; \quad j,k = 1,\ldots,m \right\} \tag{10}$$

Consider, now, the result of [21]: the information matrix of the m.l. estimators of $P = (\omega,\delta)$ in the model (1) is

---

[5] Integrable in the sense of §I and bounded on $[0,2\pi]$.

$$I_p = \left\{ \frac{1}{2\pi} \frac{\sigma_x^2}{\sigma_e^2} \int_0^{2\pi} \frac{\partial G}{\partial p_j}(\theta) \frac{\partial \bar{G}}{\partial p_k}(\theta) \frac{f_x(\theta)}{f_e(\theta)} d\theta;\ j,k = 1,\ldots,m \right\} \tag{11}$$

where

$$G(\theta) = \frac{\omega(\theta)}{\delta(\theta)} \tag{12}$$

and

$$\omega(\theta) = \omega_0 + \omega_1 e^{-i\theta} + \ldots + \omega_p e^{-pi\theta} \tag{13}$$

$$\delta(\theta) = 1 - \delta_1 e^{-i\theta} - \ldots - \delta_q e^{-qi\theta} \tag{14}$$

The analogy with the standard set-up presented above (4) suggests that one considers the model

$$z_\theta = \frac{\partial G}{\partial p_i}(\theta)\, p_1 + \ldots + \frac{\partial G}{\partial p_m}(\theta)\, p_m + \varepsilon_\theta \tag{15}$$

where z_θ, ε_θ are complex r.v.,

$\varepsilon_\theta = \xi_\theta + i\,\eta_\theta$, ξ_θ, η_θ being independent real r.v. with mean 0 and variances

$$\mathrm{Var}(\xi_\theta) = \mathrm{Var}(\eta_\theta) = 2\pi \frac{\sigma_e^2}{\sigma_x^2} f_e(\theta) \tag{16}$$

and such that the ξ's are independent for different θ's.

Proof:

Let $z = x + iy$, $G = G_1 + iG_2$, then (15) splits into two parts:

$$\begin{cases} x_\theta = \frac{\partial G_1}{\partial p_1}(\theta)\, p_1 + \ldots + \frac{\partial G_1}{\partial p_m}(\theta)\, p_m + \xi_\theta & (17) \\ y_\theta = \frac{\partial G_2}{\partial p_1}(\theta)\, p_1 + \ldots + \frac{\partial G_2}{\partial p_m}(\theta)\, p_m + \eta_\theta & (18) \end{cases}$$

From (17) the information matrix for the p's is

$$I_{p,1} = \left\{ \frac{1}{2\pi} \frac{\sigma_x^2}{\sigma_e^2} \int_0^{2\pi} \frac{\partial G_1}{\partial p_j}(\theta) \frac{\partial G_1}{\partial p_k}(\theta) \frac{f_x(\theta)}{f_e(\theta)}\, d\theta \right\} \tag{19}$$

and from (18)

$$I_{p,2} = \left\{ \frac{1}{2\pi} \frac{\sigma_x^2}{\sigma_e^2} \int_0^{2\pi} \frac{\partial G_2}{\partial p_j}(\theta) \frac{\partial G_2}{\partial p_k}(\theta) \frac{f_x(\theta)}{f_e(\theta)}\, d\theta \right\} \tag{20}$$

i.e., using the additivity of the information matrices and (16)

$$I_p = \left\{ \frac{1}{2\pi} \frac{\sigma_x^2}{\sigma_e^2} \int_0^{2\pi} \left(\frac{\partial G_1}{\partial p_j} \frac{\partial G_1}{\partial p_k} + \frac{\partial G_2}{\partial p_j} \frac{\partial G_2}{\partial p_k} \right) \frac{f_x(\theta)}{f_e(\theta)}\, d\theta \right\} \tag{21}$$

and, since (13) is real (f_x and f_e are symmetric)

$$I_p = \left\{ \frac{1}{2\pi} \frac{\sigma_x^2}{\sigma_e^2} \int_0^{2\pi} \frac{\partial G}{\partial p_j}(\theta) \frac{\partial \bar{G}}{\partial p_k}(\theta) \frac{f_x(\theta)}{f_e(\theta)}\, d\theta \right\} \tag{22}$$

which is the desired result.

This provides the proof of the following:

<u>Theorem I</u>: A design $f_x(\theta)$ is D-optimal for the model (1) iff it is D-optimal for the model (15).

The main advantage of Th. I is that it provides powerful methods and results for the solution of the original problem. The characterization of the solution to be used now is completely different from the approach of Box and Jenkins: part of Minnich's work [17] seems to prove that the consideration of the autocorrelations of the x_t-process, which was very useful, rapid and elegant in a simple example, leads to hopelessly complex computations as soon as one considers more than 2 parameters.

A summary of the results on D-optimality to be used later, will be given in the next section.

At this point it is important to make some comments concerning the specific aspects of the design for dynamic models problem:

1) For dynamic models, the design problem has some favorable aspects

- The "control" variable θ is one-dimensional,
- Its range is well defined ($[0,2\pi]$),
- The solution is symmetric $f(2\pi-\theta) = f(\theta)$.

2) This is a little compensation to the main difficulty, carefully avoided until now, namely the factor σ_x^2. The introduction of σ_x^2 in (4) was completely artificial: the importance of σ_x^2 is a consequence of the initial problem (find an optimal input for the model (1)). An obvious solution to "improve" a design would be to increase σ_x^2: this is of course completely unrealistic since, in most real life situations, the model (1) can be considered as a good representation of the system under study only in a neighborhood of

the mean values (taken here for convenience and without loss of generality, equal to zero) of x_t and y_t. It is therefore necessary to impose some constraint on the ranges of x_t and/or y_t: this is done by limiting the corresponding variances. Three types of constraints were considered by Box and Jenkins [3]:

$$C_1: \ \sigma_x^2 \leq c_1 \tag{23}$$

$$C_2: \ \sigma_y^2 \leq c_2 \tag{24}$$

$$C_3: \ \sigma_x^2 \sigma_y^2 \leq c_3 \tag{25}$$

It turns out that C_1 and C_2 are special cases of the so-called linear constraint:

$$C_L: \ a\,\sigma_x^2 + b\,\sigma_t^2 \leq c_L \qquad a,b \geq 0 \tag{26}$$

considered by McGregor [16] in relation to some control theory problems. The quantity one is interested in maximizing is now better written as

$$(\sigma_x^2)^m \ \mathrm{Det}(I_{ED}) \tag{27}$$

where in general both I_{ED} (the information matrix of the experimental design theory) and σ_x^2 are functions of $f_x(\theta)$ since

$$\sigma_y^2 = \sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 \, f_x(\theta) d\theta + \sigma_e^2 \tag{28}$$

where

$$G(\theta) = \omega(e^{-i\theta})/\delta(e^{-i\theta}) \tag{29}$$

[See Jenkins and Watts [12] for a proof of (28)]. Under the constraint C_1, the two parts of (28) are independent: the maximum is reached by taking $\sigma_x^2 = c_1$ and maximizing $Det(I_{ED})$ without constraint, i.e. the results of the classical theory of experimental design apply directly. In the other cases, it is apparent that the optimum implies that the inequality in C_2, C_3 or C_L is in fact an equality.

This section will now be concluded with a point of terminology.

In order to avoid some confusion with the terminology "discrete" and "continuous" used in the literature of experimental designs in relation with the possible values of the control variable θ, a solution (i.e. a design) will be called "stochastic" if $f_{x;2} = 0$[6], "deterministic" if $f_{x;1} = 0$ and mixed in all other cases. The justification for this comes from the spectral decomposition of a stationary stochastic process which shows that a jump in the spectral distribution function at θ_0 (i.e. with the convention here a $\sigma_0^2\delta_{\theta_0}$ (Dirac) in the density) corresponds to a deterministic sine-wave with random amplitude

$$\zeta \cos\theta t \tag{30}$$

[6] Recall $f_x = f_{x;1} + f_{x;2}$, $f_{x;1}$ a.e. continuous, $f_{x;2}$ discrete.

with

$$E(\zeta) = 0 \qquad Var(\zeta) = \sigma_0^2 \tag{31}$$

In the problem under consideration here, since ζ is determined once and for all, it seems realistic to use for ζ the distribution assigning the same probability to the two points $+\sigma$, $-\sigma$, and 0 elsewhere, since the stochastic process (30) is not ergodic. The solution corresponding to a discrete mass is then truly "deterministic". On the other hand, when $f_x = f_{x;1}$, the spectral decomposition

$$x_t = \int_0^{2\pi} e^{it\theta} \, d\xi(\theta) \tag{32}$$

where $\xi(\lambda)$ is a stochastic process with orthogonal increments and such that

$$Var(\xi(\theta)) = f_x(\theta)d\theta \tag{33}$$

shows that x_t is the "sum" of many elementary and mutually orthogonal harmonic oscillations

$$e^{it\theta} \, d\xi(\theta) \tag{34}$$

i.e. corresponds really to the idea of a stochastic process. Of course, the stochastic solutions in which one is primarily interested here are the ARMA (autoregressive-moving average) processes i.e. stochastic processes with rational spectral densities and a small number of parameters.

III. Some Important Results on D-Optimality

III.1. Properties of information matrices.

The following theorem, proved in Fedorov [8] summarizes the properties of information matrices:

P1) For any design f_x, the information matrix $I(f_x)$ is symmetric, semi-definite and positive,

P2) The information matrix is singular if the number of different points of the design is less than the number m of parameters,

P3) The family of matrices $I(f_x)$ corresponding to all possible normalized[7] designs is convex,

P4) For any design f_x, the matrix $I(f_x)$ can be represented as

$$I(f_x) = \left\{ \frac{1}{2\pi} \frac{\sigma_x^2}{\sigma_e^2} \sum_{\nu=1}^{n} \frac{\partial G}{\partial p_j}(\theta_\nu) \frac{\partial \bar{G}}{\partial p_k}(\theta_\nu) \frac{p_\nu}{f_e(\theta_\nu)} \; ; \; j,k = 1,\ldots,m \right\} \tag{35}$$

$$\text{with} \quad \begin{cases} n \leq m(m+1)/2 + 1 \\ 0 \leq p_\nu \leq 1 \qquad \sum_{\nu=1}^{n} p_\nu = 1 \end{cases} \tag{36}$$

P4 is certainly the most useful property here: it shows that one can restrict the research for D-optimal designs to deterministic designs with a finite number of points. This property will be widely used for the characterization and construction of optimal designs.

---

[7] i.e., here, with the same σ_x^2.

III.2. The Equivalence Theorem and its consequences.

It is a common practice in the theory of experimental designs to consider the quantity

$$d_{f_x}(\theta) = \mathrm{Var}(\hat{z}_\theta)/\mathrm{Var}(\varepsilon_\theta) \tag{37}$$

where $\hat{z}_\theta$ represent the predicted value at θ when the parameters of (15) have been estimated using the design $f_x(\theta)$. In the general case this is a real function of a multi-dimensional variable and the situation here is very favorable since θ is real. One has

$$\mathrm{Var}(\hat{z}_\theta) = \left[\frac{\partial G}{\partial P}(\theta)\right]\left[I_P(f_x)\right]^{-1}\left[\frac{\partial G}{\partial P}(\theta)\right]^* \tag{38}$$

with

$$\left[\frac{\partial G}{\partial P}(\theta)\right] = \left[\frac{\partial G}{\partial p_1}(\theta) \;\ldots\; \frac{\partial G}{\partial p_m}(\theta)\right] \tag{38'}$$

Defining now

$$I_P(\delta_\theta) = \left\{ \frac{1}{4\pi}\frac{\sigma_x^2}{\sigma_e^2}\left\{\left[\frac{\partial G}{\partial P}(\theta)\right]\left[\frac{\partial G}{\partial P}(\theta)\right]^* + \left[\frac{\partial G}{\partial P}(2\pi-\theta)\right]\left[\frac{\partial G}{\partial P}(2\pi-\theta)\right]^*\right\}\frac{1}{f_e(\theta)}\right\} \tag{39}$$

one has

$$I_P(f_x) = \int_0^{2\pi} I_P(\delta_\theta)\, f_x(\theta)d\theta \tag{40}$$

and, as soon as f_x is not a degenerate design

$$d_{f_x}(\theta) = \mathrm{Tr}\left\{[I_P(f_x)]^{-1}\, I_P(\delta_\theta)\right\} \tag{41}$$

Now, considering the explicit form of the partial derivatives of $G(\theta)$, it is possible to derive the expression:

$$d_{f_x}(\theta) = \frac{1}{|\delta(\theta)|^4 \, f_e(\theta)} \, Q_m(\cos\theta, \sin\theta) \tag{42}$$

where $Q_m(\cos\theta, \sin\theta)$ is a polynomial of degree $(m-1)$ in $\cos\theta, \sin\theta$. This implies that $d_{f_x}(\theta)$ is continuous iff $f_e(\theta)$ is continuous. Using (40) and (41) it is easy to check that

$$\int_0^{2\pi} d_{f_x}(\theta) \, f_x(\theta) d\theta = m \tag{43}$$

for all $f_x(\theta)$ such that the design is not degenerate. Note that, because of the presence of $f_x(\theta)$, the integral over $[0,2\pi]$ in (43) is in fact the sum over the points of the design.

The Equivalence Theorem (Karlin and Studden [13]) says that the three conditions are equivalent:

i) The design f_x^* is D-optimal,

ii) f_x^* minimizes $\max\limits_{\theta} d_{f_x}(\theta)$

iii) $d_{f_x^*}(\theta) \leq m$ for all $\theta \varepsilon [0,2\pi]$.

It is not difficult to check, using (43) and condition iii), that for an optimal design f_x^*, $d_{f_x^*}(\theta) = m$ at the points of the design.

The characterization of an optimal design as a design such that for all $\theta \varepsilon [0,2\pi]$

$$\Delta^1_{f_x}(\theta) = d_{f_x}(\theta) - m \leq 0 \tag{44}$$

will be extensively used (after suitable modification) in the following: $\Delta^1_{f_x}(\theta)$ is not only one mean to express the optimality of a design, but is also a pivotal quantity in the derivation of an iterative method of construction of optimal designs.

Given f_X non degenerate, consider the design

$$f_{X;\alpha} = (1-\alpha)f_X + \alpha\,\delta_\theta \qquad \alpha\varepsilon[0,1) \tag{45}$$

then

$$I_p(f_{X;\alpha}) = (1-\alpha)I_p(f_X) + \alpha\, I_p(\delta_\theta) \tag{46}$$

and

$$\frac{d}{d\alpha} I_p(f_{X;\alpha}) = I_p(\delta_\theta) - I_p(f_X) \tag{47}$$

Now it is well known (Fedorov [8]), that

$$\frac{d}{d\alpha} \text{Log Det}(M(\alpha)) = \text{Tr}\left\{M^{-1}\frac{dM}{d\alpha}\right\} \tag{48}$$

giving here

$$\frac{d}{d\alpha} \text{Log Det}(I_p(f_{X;\alpha}))\Big|_{\alpha=0} = \text{Tr}\left\{[I_p(f_X)]^{-1}[I_p(\delta_\theta) - I_p(f_X)]\right\} \tag{49}$$

$$= d_{f_X}(\theta) - m \tag{50}$$

$$= \Delta^1_{f_X}(\theta) \tag{51}$$

In addition:

$$\frac{d^2}{d\alpha^2} \text{Log Det}(I_p(f_{X;\alpha}))\Big|_{\alpha=0} = -\text{Tr}\left\{[I_p(f_X)]^{-1}[I_p(\delta_\theta) - I_p(f_X)]\right.$$
$$\left.[I_p(f_X)]^{-1}[I_p(\delta_\theta) - I_p(f_X)]\right\} \tag{52}$$

Let

$$A = [I_p(f_x)]^{-1}[I_p(\delta_\theta) - I_p(f_x)] \tag{53}$$

there exists an orthogonal matrix B such that

$$A = B\Lambda B' \tag{54}$$

where Λ is a diagonal matrix. Now, using a property of the Trace operator

$$Tr(A_1A_2) = Tr(A_2A_1) \tag{55}$$

$$Tr(AA) = Tr(\Lambda B'B\Lambda B'B) = Tr(B\Lambda B'B\Lambda B') \tag{56}$$

$$= Tr(\Lambda\Lambda) \geq 0 \tag{57}$$

so that

$$\frac{d^2}{d\alpha^2} \text{Log Det } I(f_{x;\alpha})\bigg|_{\alpha=0} \leq 0 \tag{58}$$

An iterative method of construction of D-optimal designs now proceeds as follows:

Step 1. Using f_0, derive $\Delta^1_{f_0}(\theta)$ and find θ_0 such that $\Delta^1_{f_0}(\theta_0)$ is maximum,
[If $\Delta^1_{f_0}(\theta_0) \leq 0$, then f_0 is an optimal solution].

Step 2. Define $f_{0;\alpha} = (1-\alpha)f_0 + \alpha\, \delta_{\theta_0}$ and find α_0 such that $\mathrm{Det}(I_m(f_{0;\alpha_0}))$ is maximum,

Step 3. Take $f_1 = (1-\alpha_0)f_0 + \alpha_0\, \delta_{\theta_0}$, replace f_0 by f_1 and go to step 1.

This procedure is completely discussed in Fedorov [8], where its convergence is proved: some elements of this proof, like the recurrence formula relating $\mathrm{Det}(I_p(f_{n+1}))$ to $\mathrm{Det}(I_p(f_n))$ will be used later. Improvements of the rate of convergence are discussed in Atwood [1]. It is to be noted that this procedure (or similar ones to be introduced later) is one of the main reasons for considering mixed designs. (Another reason being, of course, the property P4 of the information matrix.)

III.3. Kiefer's inequality.

Using (51) and (58) for f_x^* D-optimal and θ fixed, Kiefer [14] derived the lower bound

$$\frac{\mathrm{Det}\ (I_p(f_x))}{\mathrm{Det}\ (I_p(f_x^*))} \geq \exp\ (-\max_{\theta} \Delta_{f_x}^1\ (\theta)) \tag{59}$$

This inequality, expressing how close a design f_x is to D-optimality using $\Delta_{f_x}^1\ (\theta)$, is another advantage of $\Delta_{f_x}^1\ (\theta)$. It is very useful for iterative methods, in order to stop the computations when a desired accuracy is reached.

III.4. Conclusions concerning the dynamic model designs.

At this point it may be useful, in the light of the results just presented, to underline the advantages as well as the limits of the relationship established in the preceding sections.

The main advantage is in the function $\Delta^{1}_{f_x}(\theta)$, which gives not only a convenient characterization of the D-optimal designs, but also the starting point for the improvement of a given design. The limits are of two kinds:

- As previously mentioned, the situation of the classical theory of experimental design corresponds to the constraint C_1 ($\sigma_x^2 \leq c_1$) for the dynamic model problem: it is necessary to extend the methods to other useful situations of constrained D-optimality.
- Many results of the classical theory are not available for dynamic models because of the form of the model (15) (special form of polynomial regression) or, when (15) is split into two trigonometric regressions, because of the form of the error variance ($\propto |\delta(\theta)|^4 f_e(\theta)$). More work in this direction is certainly needed (see Karlin and Studden [13]).

IV. Constrained D-Optimality

It was mentioned at the end of Section II that the notion of D-optimality for the design problem corresponding to the dynamic model (1) was more involved than the corresponding notion in experimental design situations, since there are some constraints on the input and/or output variances to be taken into account. The influence of these constraints will be investigated now.

The quantity to be maximized is

$$\mathrm{Det}(I_p(f_x)) = (\sigma_x^2)^m \, D(f_x) \tag{60}$$

where, by definition

$$D(f_x) = \text{Det}\left(\left\{\frac{1}{2\pi\,\sigma_e^2}\int_0^{2\pi} \frac{\partial G}{\partial p_j}(\theta)\,\frac{\partial \bar{G}}{\partial p_k}(\theta)\,\frac{f_x(\theta)}{f_e(\theta)}\,d\theta\right\}\right) \tag{61}$$

(60) shows clearly that, under the constraint C_1 $(\sigma_x^2 \leq c_1)$, σ_x^2 should be chosen as large as possible and independently $D(f_x)$ made as large as possible: this is exactly the situation of Section III.

Since all the other realistic constraints use the output variance σ_y^2, it is now necessary to consider the general model. Recalling the condition of independence between $\{x_t\}$ and $\{e_t\}$, one has

$$\sigma_y^2 = \sigma_x^2 \int_0^{2\pi} |G(\theta)|^2\, f_x(\theta)d\theta + \sigma_e^2 \tag{62}$$

where

$$G(\theta) = \frac{\omega(e^{-i\theta})}{\delta(e^{-i\theta})} \tag{63}$$

[8]

As a consequence of the constraints C_2, C_3, σ_x^2 is to be considered as a function of the solution $f_x(\theta)$.

---

[8] Because of the assumptions made concerning the roots of $\omega(z)$ and $\delta(z)$, $|G(\theta)|$ which is a continuous function of θ, is bounded not only above, but also away from 0.

IV.1. Constraint C_2: $\sigma_y^2 \leq c_2$.

It is first obvious that an optimal design will have $\sigma_y^2 = c_2$: for this reason, σ_y^2 will be consistently used instead of c_2.

Consider then the equivalent problem:

$$\begin{cases} \text{maximize} \quad L(f_x) = m \operatorname{Log} \sigma_x^2 + \operatorname{Log} D(f_x) \\ \text{subject to} \quad \sigma_x^2 = \dfrac{\sigma_y^2 - \sigma_e^2}{\int_0^{2\pi} |G(\theta)|^2 f(\theta) d\theta} \end{cases} \tag{64}$$

(of course it is assumed that $\sigma_y^2 > \sigma_e^2$!)

Take any non-degenerate design f_1, and any $f_2 \neq f_1$ and consider the (non-degenerate)

$$f_\alpha = (1-\alpha) f_1 + \alpha f_2 \qquad \alpha \varepsilon [0,1) \tag{65}$$

then

$$L(\alpha) = L(f_\alpha) = m \operatorname{Log} \sigma_x^2 + \operatorname{Log} D(f_\alpha) \tag{66}$$

and, because of (47),(48) and (64)

$$\frac{d}{d\alpha} L(\alpha) = -m \frac{\int_0^{2\pi} |G(\theta)|^2 (f_2(\theta) - f_1(\theta)) d\theta}{\int_0^{2\pi} |G(\theta)|^2 f_\alpha(\theta) d\theta} + \operatorname{Tr}\left\{[I(f_\alpha)]^{-1} [I(f_2) - I(f_1)]\right\} \tag{67}$$

and

$$\frac{d^2}{d\alpha^2} L(\alpha) = m\left[\frac{\int_0^{2\pi} |G(\theta)|^2(f_2(\theta)-f_1(\theta))d\theta}{\int_0^{2\pi} |G(\theta)|^2 f_\alpha(\theta)d\theta}\right]^2 - Tr\left\{[I(f_\alpha)]^{-1}[I(f_2)-I(f_1)] [I(f_\alpha)]^{-1}[I(f_2)-I(f_1)]\right\} \tag{68}$$

A necessary condition for f_1 to be optimal is

$$\frac{d}{d\alpha} L(\alpha)\Big|_{\alpha=0} \leq 0 \qquad \text{for all } f_2 \tag{69}$$

i.e.

$$Tr\left\{[I(f_1)]^{-1}[I(f_2)-I(f_1)]\right\} - m\frac{\int_0^{2\pi} |G(\theta)|^2(f_2(\theta)-f_1(\theta))d\theta}{\int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta} \leq 0 \tag{70}$$

or

$$Tr\left\{[I(f_1)]^{-1} I(f_2)]\right\} - m\frac{\int_0^{2\pi} |G(\theta)|^2 f_2(\theta)d\theta}{\int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta} \leq 0 \tag{71}$$

In order to prove that f_1 satisfying (69) is in fact optimal, it is sufficient to show that

$$\frac{d^2}{d\alpha^2} L(\alpha)\Big|_{\alpha=0} < 0 \tag{72}$$

for all f_2 such that

$$\frac{d}{d\alpha} L(\alpha)\Big|_{\alpha=0} = 0 \tag{73}$$

Let λ_ν, $\nu = 1,\ldots,m$ be the eigenvalues of $[I(f_1)]^{-1}[I(f_2)]$. Then (71) and (73) imply that

$$\frac{\int_0^{2\pi} |G(\theta)|^2 f_2(\theta)d\theta}{\int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta} = \frac{\sum_{\nu=1}^{m} \lambda_\nu}{m} = \bar{\lambda} \tag{74}$$

and with this notation, and using an argument similar to (53)-(57)

$$\frac{d^2}{d\alpha^2} L(\alpha) = m\,\bar{\lambda}^2 - \sum_{\nu=1}^{m} \lambda_\nu^2 \leq 0 \tag{75}$$

[The only possible case of equality in (75) is when all the λ's are equal. This implies that

$$[I(f_1)]^{-1} I(f_2) = k\, II_m \tag{76}$$

(II_m is the $m\times m$ identity matrix), or since $k=1$ the designs being normalized,

$$I(f_1) = I(f_2) \tag{77}$$

i.e. one has equality in (75) iff f_2 is D-optimal.]

These results are summarized in the following Theorem:

Theorem II. The design f_x^* is D-optimal iff, for all f_x

$$\forall f_x: \ \mathrm{Tr}\left\{[I(f_x^*)]^{-1}\ I(f_x)]\right\} \leq m \frac{\int_0^{2\pi} |G(\theta)|^2 f_x(\theta)d\theta}{\int_0^{2\pi} |G(\theta)|^2 f_x^*(\theta)d\theta} \tag{78}$$

Recalling now the property P4 of the information matrix, the design f_x^* is equivalent to a discrete design. To express the fact that it can not be improved, it is sufficient to take for f_x any design "on one point"[9] : δ_θ for instance.

Defining (see (37) and (41))

$$\Delta^2_{f_x}(\theta) = d_{f_x}(\theta) - m \frac{|G(\theta)|^2}{\int_0^{2\pi} |G(\theta)|^2 f_x(\theta)d\theta} \tag{79}$$

it is possible to give the modified version of the equivalence theorem.

Theorem III. The three conditions are equivalent:

i) The design f_x^* is D-optimal for the constrained problem (62)

ii) f_x^* minimizes $\max\limits_{\theta} \Delta^2_{f_x}(\theta)$.

iii) $\Delta^2_{f_x^*}(\theta) \leq 0$ for all $\theta\varepsilon[0,2\pi]$.

Proof.

i) and iii) have already been shown to be equivalent. Now one has

$$\int_0^{2\pi} \Delta^2_{f_x}(\theta)\ f_x(\theta)d\theta = 0 \tag{80}$$

---

[9] "On one point" means in fact "on two points symmetric with respect to π."

for all non degenerate design f_x. This implies, since $f_x(\theta) \geq 0$, that

$$\max_{\theta} \Delta^2_{f_x}(\theta) \geq 0 \tag{81}$$

So, if f_x^* is D-optimal, iii) implies that (81) is minimum. Conversely, if (81) is minimum, then iii) is verified and the design is D-optimal.

It will now be shown how an iterative method of construction of a sequence of designs, converging to a D-optimal design, proceeds along the same lines as in the classical theory, with suitable modifications. Consider the particular case of (65):

$$f_{x;\alpha} = (1-\alpha)f_x + \alpha\, \delta_{\theta} \tag{82}$$

then (67) gives:

$$\frac{d}{d\alpha} L(\alpha)\Big|_{\alpha=0} = \Delta^2_{f_x}(\theta) \tag{83}$$

i.e., if one wants to improve the design f_x by addition of one point, the "best" choice is $\tilde{\theta}$ such that

$$\tilde{\theta} \text{ maximizes } \Delta^2_{f_x}(\theta) \tag{84}$$

[If f_x is not D-optimal, $\max \Delta^2_{f_x}(\theta)$ is positive, see (80).] The best value of α is $\tilde{\alpha}$ such that

$$\tilde{\alpha} \text{ maximizes } L(\alpha) \qquad \alpha\varepsilon[0,1) \tag{85}$$

The iterative method is as follows:

Step 1. Using f_0, derive $\Delta^2_{f_0}(\theta)$, and find θ_0 such that $\Delta^2_{f_0}(\theta_0)$ is maximum.

[If $\Delta^2_{f_0}(\theta_0)$ is equal to zero, then f_0 is an optimal solution for the constrained problem.]

Step 2. Define $f_{0;\alpha} = (1-\alpha)f_0 + \alpha\ \delta_{\theta_0}$ and find α_0 such that $L(\alpha_0)$ is maximum.

Step 3. Take $f_1 = (1-\alpha_0)f_0 + \alpha_0\ \delta_{\theta_0}$, replace f_0 by f_1 and return to step 1.

The convergence of this procedure is stated as

<u>Theorem IV</u>: The above procedure is convergent.

The proof of the convergence is somewhat more involved than in the preceding case (see Fedorov [8]). It rests on the fact that since f_0 is non-degenerate

$$0 < \text{Det}\ (I_p(f_0)) < \text{Det}\ (I_p(f_n)) < \text{Det}(I_p(f^*)) < \infty \tag{86}$$

the sequence $\{\text{Det}\ (I_p(f_n))\}$ being by construction strictly increasing.

<u>Lemma</u>: The function $d_{f_n}(\theta)$ is uniformly bounded

$$\exists M > 0: \ \forall n\ , \forall\theta \qquad 0 \leq d_{f_n}(\theta) < M \tag{87}$$

Proof of the Lemma.

$$d_{f_n}(\theta) = \mathrm{Tr}\left\{[I_p(f_n)]^{-1}\, I_p(\delta_\theta)\right\} \tag{88}$$

Let $\lambda_1^{(n)},\ldots,\lambda_m^{(n)}$ be the eigenvalues of $I_p(f_n)$. All the elements of $I_p(f_n)$ are uniformly bounded in modulus by the quantity

$$Q = \left\{\max_\theta \frac{\sigma_y^2 - \sigma_e^2}{|G(\theta)|^2}\right\} \frac{1}{2\pi\,\sigma_e^2} \left\{\max_\theta \frac{Q_m(\sin\theta,\cos\theta)}{|\delta(\theta)|^4\; f_e(\theta)}\right\} < +\infty \tag{89}$$

and then the eigenvalues $\lambda_\nu^{(n)}$ are uniformly bounded above by $\Lambda = mQ$

$$\lambda_\nu^{(n)} \leq \Lambda < +\infty \tag{90}$$

As a consequence of (86) and (90) there exists a uniform lower bound λ $(\leq \Lambda^{-m})$ such that

$$0 < \lambda < \lambda_\nu^{(n)} \qquad \text{for all } n \tag{91}$$

Now Q, defined in (89) is also a bound for the modulus of the elements of $I_p(\delta_\theta)$.

In addition, there exists an orthogonal matrix P_n diagonalizing $[I_p(f_n)]$ then

$$d_{f_n}(\theta) = \mathrm{Tr}\left\{\begin{bmatrix} \frac{1}{\lambda_1^{(n)}} & & \\ & \ddots & \\ & & \frac{1}{\lambda_m^{(n)}} \end{bmatrix} P_n\, I_p(\delta_\theta)\, P_n'\right\} \tag{92}$$

and, since the elements of P_n are uniformly bounded by 1,

$$d_{f_n}(\theta) \leq m\,\Lambda^m\,Q = M < +\infty \tag{93}$$

Note that this lemma proves an intuitive fact, and excludes the existence of pathological sequences of designs where

$$\max_\theta d_{f_n}(\theta) \to +\infty \quad \text{as} \quad \mathrm{Det}\,(I_m(f_n)) \text{ increases.}$$

It is now possible to proceed to the proof of the theorem:

Proof of Theorem IV.

It was shown in Fedorov [8] that

$$f_{n;\alpha} = (1-\alpha)f_n + \alpha\,\delta_{\theta_n} \tag{94}$$

implies

$$D(f_{n;\alpha}) = (1-\alpha)^m \left\{1 + \frac{\alpha}{1-\alpha} d_{f_n}(\theta_n)\right\} D(f_n) \tag{95}$$

in addition, corresponding to $f_{n;\alpha}$ one has

$$(\sigma_x^2)_{n;\alpha} = \frac{\sigma_y^2 - \sigma_e^2}{\int_0^{2\pi} |G(\theta)|^2((1-\alpha)f_n(\theta)+\alpha\,\delta_{\theta_n})d\theta} \tag{96}$$

and then

$$\text{Det}\,(I_p(f_{n;\alpha})) = \frac{1 + \frac{\alpha}{1-\alpha}\, d_{f_n}(\theta_n)}{\left[1 + \frac{\alpha}{1-\alpha} \frac{|G(\theta_n)|^2}{\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta}\right]^m}\, \text{Det}\,(I_p(f_n)) \tag{97}$$

$$= R_n(\alpha)\, \text{Det}\,(I_p(f_n)) \tag{98}$$

The iterative method outlined above consists in choosing α so that $R_n(\alpha)$ is as large as possible, i.e. α_n is solution of

$$\frac{d}{d\alpha} R_n(\alpha)\Big|_{\alpha=\alpha_n} = 0 \tag{99}$$

giving

$$\alpha_n = \frac{\Delta^2_{f_n}(\theta_n)}{\Delta^2_{f_n}(\theta_n) + (m-1)d_{f_n}(\theta_n)\, \frac{|G(\theta_n)|^2}{\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta}} \tag{100}$$

If $\Delta^2_{f_n}(\theta_n) \to 0$ as $n \to +\infty$, then by Theorem III, $f_n \to f^*$.

The assumption that the method does not converge to f^* is equivalent to the assumption that there exists $K>0$ such that

$$\Delta^2_{f_n}(\theta_n) > K \text{ for infinitely many } n \tag{101}$$

Because of the lemma and the assumptions on $\omega(z)$ and $\delta(z)$, there exists $K' < +\infty$ such that for all n

$$d_{f_n}(\theta_n) \frac{|G(\theta_n)|^2}{\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta} < K' \tag{102}$$

and one has then

$$\alpha_n > \frac{K}{K + K'} \text{ for infinitely many } n \tag{103}$$

i.e.

$$\alpha_n \not\to 0 \qquad n \to \infty \tag{104}$$

The fact that the sequence Det $(I_p(f_n))$ is increasing (strictly) and bounded by Det $(I_p(f^*))$ will now be used to show that the assumption $\Delta^2_{f_n}(\theta_n) \not\to 0$ leads to a contradiction.

For all n, choose β_n

$$0 < \beta_n \leq \alpha_n \tag{105}$$

β_n small enough to justify the approximation:

$$\frac{1 + \frac{\beta_n}{1-\beta_n} d_{f_n}(\theta_n)}{\left[1 + \frac{\beta_n}{1-\beta_n} \frac{|G(\theta_n)|^2}{\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta}\right]^m} \simeq 1 + \frac{\beta_n}{1-\beta_n}\left(d_{f_n}(\theta_n) - m \frac{|G(\theta_n)|^2}{\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta}\right) \tag{106}$$

$$\simeq 1 + \frac{\beta_n}{1-\beta_n} \Delta^2_{f_n}(\theta_n) \tag{107}$$

since for infinitely many n

$$\alpha_n > \frac{K}{K + K'} > 0 \tag{108}$$

there exists $\beta>0$ such that, for those n,

$$0 < \beta < \frac{\beta_n}{1-\beta_n} < \alpha_n \tag{109}$$

and then

$$R_n(\alpha_n) \geq R_n(\beta_n) > 1 + \beta K \tag{110}$$

Now for all n

$$R_n(\alpha_n) \geq 1 \tag{111}$$

and because of (110) the infinite product

$$\prod_{n=1}^{\infty} R_n(\alpha_n) \quad \text{is divergent} \tag{112}$$

i.e.

$$\text{Det}\,(I_m(f_n)) \to \infty \qquad n \to \infty \tag{113}$$

in opposition with the fact that $\text{Det}\,(I_m(f_n))$ is bounded.

In practical applications the iterative procedure will be stopped as soon as

$$\Delta^2_{f_n}(\theta_n) = d_{f_n}(\theta_n) - m \frac{|G(\theta_n)|^2}{\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta} < \varepsilon \tag{114}$$

where ε is determined by the modified version of Kiefer's inequality, valid in the neighborhood of an optimal design:

<u>Lemma</u>: Let f_x be in the neighborhood of an optimal design f_x^*, then

$$\frac{\text{Det}\,(I_p(f_x))}{\text{Det}\,(I_p(f_x^*))} \geq \exp\left\{- \max_\theta \Delta^2_{f_x}(\theta)\right\} \tag{115}$$

Proof:

Consider

$$f_\alpha = \alpha\, f_x^* + (1-\alpha)f_x \tag{116}$$

Then

$$\frac{d}{d\alpha} \text{Log Det}\,(I_p(f_\alpha))\Big|_{\alpha=0} = \text{Tr}\left\{[I_p(f_x)]^{-1}\, I_p(f^*)\right\} - m\, \frac{\int_0^{2\pi} |G(\theta)|^2 f_x^*(\theta)d\theta}{\int_0^{2\pi} |G(\theta)|^2 f_x(\theta)d\theta} \tag{117}$$

and, since f_x^* is equivalent to a discrete design

$$\frac{d}{d\alpha} \text{Log Det } (I_p(f_\alpha))\Big|_{\alpha=0} \le \max_\theta \Delta^2_{f_x}(\theta) \tag{118}$$

Now for $\max_\theta \Delta^2_{f_x}(\theta)$ small enough, the proof of (75) shows that

$$\frac{d^2}{d\alpha^2} \text{Log Det } (I_p(f_\alpha)) \le 0 \tag{119}$$

i.e.

$$\frac{d}{d\alpha} \text{Log Det } (I_p(f_\alpha)) \le \max \Delta^2_{f_x}(\theta) \qquad \alpha\varepsilon[0,1] \tag{120}$$

and

$$\frac{\text{Det } (I_p(f_x))}{\text{Det } (I_p(f^*_x))} \ge \exp\left\{- \max_\theta \Delta^2_{f_x}(\theta)\right\} \tag{121}$$

IV.2. Linear constraint C_L: $a\,\sigma^2_x + b\,\sigma^2_y \le c_L$ $(a,b \ge 0)$.

C_1, C_2 are clearly special cases of the linear constraint on σ^2_x, σ^2_y. Only the general form of the equivalence theorem will be given here, the proofs being exactly as in IV.1.

C_L is to be expressed as

$$\sigma^2_x \left\{a + b \int_0^{2\pi} |G(\theta)|^2\, f_x(\theta)d\theta\right\} + b\,\sigma^2_e = c_L \tag{122}$$

Considering any two designs f_1, f_2 (f_1 non degenerate) and

$$f_\alpha = (1-\alpha)f_1 + \alpha f_2 \tag{123}$$

one has:

$$\frac{d}{d\alpha} \text{Log } (\sigma_x^2)_\alpha = - \frac{b \int_0^{2\pi} |G(\theta)|^2 (f_2(\theta) - f_1(\theta))d\theta}{b \int_0^{2\pi} |G(\theta)|^2 f_\alpha(\theta)d\theta + a} \tag{124}$$

$$\frac{d^2}{d\alpha^2} \text{Log } (\sigma_x^2)_\alpha = \left[\frac{b \int_0^{2\pi} |G(\theta)|^2 (f_2(\theta) - f_1(\theta))d\theta}{b \int_0^{2\pi} |G(\theta)|^2 f_\alpha(\theta)d\theta + a} \right]^2 \tag{124}$$

and

$$\frac{d}{d\alpha} \text{Log Det } (I_p(f_\alpha))\Big|_{\alpha=0} = -m \frac{b \int_0^{2\pi} |G(\theta)|^2 (f_2(\theta) - f_1(\theta))d\theta}{b \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta + a}$$

$$+ \text{Tr} \left\{ [I_p(f_1)]^{-1} [I_p(f_2) - I_p(f_1)] \right\} \tag{125}$$

$$\frac{d^2}{d\alpha^2} \text{Log Det } (I_p(f_\alpha))\Bigg|_{\alpha=0} = m\left[\frac{b\int_0^{2\pi} |G(\theta)|^2(f_2(\theta) - f_1(\theta))d\theta}{b\int_0^{2\pi} |G(\theta)|^2 \, f_1(\theta)d\theta + a}\right]^2$$

$$- \text{Tr}\left\{[I_p(f_1)]^{-1}[I_p(f_2) - I_p(f_1)]\right.$$

$$\left.[I_p(f_1)]^{-1}[I_p(f_2) - I_p(f_1)]\right\} \quad (126)$$

Defining:

$$\Delta^L_{f_x}(\theta) = d_{f_x}(\theta) - m\,\frac{b|G(\theta)|^2 + a}{b\int_0^{2\pi} |G(\theta)|^2 \, f_x(\theta)d\theta + a} \quad (127)$$

it is possible to state directly the

Theorem V. The three conditions are equivalent:

i) The design f^*_x is D-optimal for C_L ,

ii) f^*_x minimizes $\max_\theta \Delta^L_{f_x}(\theta)$,

iii) $\Delta^L_{f_x}(\theta) \leq 0$ for all $\theta\varepsilon[0,2\pi]$.

The generalization of the iterative method for construction of C_L-constrained D-optimal designs, the proof of its convergence and the proof of the validity of Kiefer's inequality are straightforward.

V.3. Constraint C_3: $\sigma_x^2\sigma_y^2 \leq c_3$.

The method used here is the same as in the preceding cases, and leads to the same type of results, but with more involved computations to prove that the necessary condition for a maximum is also sufficient. The convergence of the corresponding iterative method for construction of C_3-constrained D-optimal designs is not proved.

As usual, consider any two designs f_1, f_2 (f_1 non degenerate) and take

$$f_\alpha = (1-\alpha)f_1 + \alpha f_2 \qquad \alpha\varepsilon[0,1) \tag{128}$$

The constraint C_4 implies:

$$(\sigma_x^2)^2 \int_0^{2\pi} |G(\theta)|^2 f_\alpha(\theta)d\theta + \sigma_x^2 \sigma_e^2 = c_3 \tag{129}$$

and

$$\frac{d}{d\alpha} \text{Log } \sigma_x^2 = - \frac{\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 (f_2(\theta) - f_1(\theta))d\theta}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_\alpha(\theta)d\theta + \sigma_e^2} \tag{130}$$

$$\frac{d^2}{d\alpha^2} \text{Log } \sigma_x^2 = \left[\frac{\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2(f_2(\theta)-f_1(\theta))d\theta}{\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_\alpha(\theta)d\theta+\sigma_e^2}\right]^2 \left[2 + \frac{\sigma_e^2}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_\alpha(\theta)d\theta+\sigma_e^2}\right] \tag{131}$$

leading to:

$$\frac{d}{d\alpha} \text{Log Det } (I_p(f_\alpha))\bigg|_{\alpha=0} = -m \frac{\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 (f_2(\theta)-f_1(\theta))d\theta}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta + \sigma_e^2}$$

$$+ \text{Tr} \left\{[I_p(f_1)]^{-1} [I_p(f_2) - I_p(f_1)]\right\} \qquad (132)$$

$$= -m \frac{\sigma_y^2 + \int_0^{2\pi} |G(\theta)|^2 f_2(\theta)d\theta}{\sigma_y^2 + \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta} + \text{Tr} \left\{[I_p(f_1)]^{-1} [I_p(f_2)]\right\} \qquad (133)$$

and

$$\frac{d^2}{d\alpha^2} \text{Log Det } (I_p(f_\alpha))\bigg|_{\alpha=0} = m \left[\frac{\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 (f_2(\theta)-f_1(\theta))d\theta}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta + \sigma_e^2}\right]^2$$

$$\left[2 + \frac{\sigma_e^2}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta + \sigma_e^2}\right]$$

$$- \text{Tr} \left\{[I_p(f_1)]^{-1} [I_p(f_2)-I_p(f_1)] [I_p(f_1)]^{-1} [I_p(f_2)-I_p(f_1)]\right\} \qquad (134)$$

A necessary condition for a maximum at f_1 is

$$\frac{d}{d\alpha} \text{Log Det } (I_p(f_\alpha))\Big|_{\alpha=0} \leq 0 \qquad \text{for all } f_2 \tag{135}$$

Using again the property P4 of the information matrix, it is possible to restrict f_2 to designs on one point, and (133), (135) give the necessary condition

$$d_{f_1}(\theta) \leq m \frac{\sigma_y^2 + \sigma_x^2 |G(\theta)|^2}{\sigma_y^2 + \sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta} \tag{136}$$

i.e., with

$$\Delta_f^3(\theta) = d_f(\theta) - m \frac{\sigma_y^2 + \sigma_x^2 |G(\theta)|^2}{\sigma_y^2 + \sigma_x^2 \int |G(\theta)|^2 f(\theta)d\theta} \tag{137}$$

the condition

$$\Delta_{f_1}^3(\theta) \leq 0 \qquad \theta\varepsilon[0,2\pi] \tag{138}$$

Note that, as in the preceding cases

$$\int_0^{2\pi} \Delta_f^3(\theta)\, f(\theta)d\theta = 0 \qquad f \text{ non degenerate} \tag{139}$$

In order to show that f_1 satisfying (138) is really a maximum, it is sufficient to show that

$$\frac{d^2}{d\alpha^2} \text{Log Det } (I_m(f_\alpha))\Big|_{\alpha=0} < 0 \tag{140}$$

for all f_2 such that

$$\frac{d}{d\alpha} \text{Log Det } (I_p(f_\alpha)) = 0 \tag{141}$$

Consider now (134), for f_1 satisfying (136).

Let λ_ν, $\nu = 1,\ldots,m$ be the eigenvalues of $[I_p(f_1)]^{-1}\ I_p(\delta_\theta)$. Then

$$d_{f_1}(\theta) = \text{Tr}\left\{[I_p(f_1)]^{-1}\ I_p(\delta_\theta)]\right\} = \sum_1^m \lambda_\nu = m\,\bar{\lambda} \tag{142}$$

and, from (141)

$$\frac{\sigma_y^2 + \sigma_x^2|G(\theta)|^2}{\sigma_y^2 + \sigma_x^2 \int_0^{2\pi} |G(\theta)|^2\ f_1(\theta)d\theta} = \bar{\lambda} \tag{143}$$

i.e.

$$\frac{\sigma_x^2(|G(\theta)|^2 - \int_0^{2\pi} |G(\theta)|^2\ f_1(\theta)d\theta)}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2\ f_1(\theta)d\theta + \sigma_e^2} = \bar{\lambda} - 1 \tag{144}$$

and (134) can be transformed into:

$$\left.\frac{d^2}{d\alpha^2} \text{Log Det } (I_p(f_\alpha))\right|_{\alpha=0} = m(\bar{\lambda}-1)^2\left\{2 + \frac{\sigma_e^2}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta + \sigma_e^2}\right\} - \sum_{\nu=1}^m (\lambda_\nu - 1)^2 \tag{145}$$

In order to discuss the sign of (145), it is necessary now to consider more closely the λ's [this is different than the linear constraint, where the 2^{nd} derivative is easily shown to be negative at the extremum].

(1) The case of $m=1$ or 2.

If $m=1$, an optimal design corresponds to two (symmetric) points where $|\delta(\theta)|^4 f_e(\theta)$ is minimum. Any other design on two (symmetric) points is either optimal, or worse than the given design[10]: in this case the first derivative is negative and it is not necessary to check for 2^{nd} derivatives.

The same remark applies for $m=2$ since, because of the symmetry, all the so-called "one-point" designs are in fact two point designs, i.e., non degenerate.

(2) The case of $m \geq 3$.

Since

$$I_m(\delta_\theta) \propto \left[\frac{\partial G}{\partial P}(\theta)\right]\left[\frac{\partial G}{\partial P}(\theta)\right]^* \tag{146}$$

it is clear that all the eigenvalues λ_ν but one are equal to zero, and that the remaining one is equal to the trace, i.e. (145) can be written as

$$\frac{d^2}{d\alpha^2} \text{Log Det}\,(I_P(f_\alpha))\Bigg|_{\alpha=0} = m\left(\frac{d_{f_1}(\theta)}{m} - 1\right)^2\left(2 + \frac{\sigma_e^2}{2\sigma_x^2\int_0^{2\pi}|G(\theta)|^2 f_1(\theta)d\theta + \sigma_e^2}\right)$$
$$- (d_{f_1}(\theta) - 1)^2 - (m-1) \tag{147}$$

---

[10] This comes from the fact that one point designs are not degenerate.

or

$$\therefore = \frac{(d_{f_1}(\theta) - m)^2}{m}\left(2 + \frac{\sigma_e^2}{\sigma_y^2 + \sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta}\right)$$

$$- (d_{f_1}^2(\theta) - 2d_{f_1}(\theta) + m) \tag{148}$$

hence the second derivative is negative iff

$$2 + \frac{\sigma_e^2}{\sigma_y^2 + \sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f_1(\theta)d\theta} < m \frac{d_{f_1}^2(\theta) - 2d_{f_1}(\theta) + m}{(d_{f_1}(\theta) - m)^2} \tag{149}$$

Let

$$\Lambda = \Lambda(\theta) = d_{f_1}(\theta) - m \qquad \Lambda\varepsilon[-m,\infty) \tag{150}$$

$$R = m \frac{d_{f_1}^2(\theta) - 2d_{f_1}(\theta) + m}{(d_{f_1}(\theta) - m)^2} = m\left(1 + 2\,\frac{m-1}{\Lambda} + m\,\frac{m-1}{\Lambda^2}\right) \tag{151}$$

then

$$\frac{dR}{d\Lambda} = -2m(m-1)\left(\frac{1}{\Lambda^2} + \frac{m}{\Lambda^3}\right)$$

Since $m \geq 3$, the minimum of $R(\Lambda)$ is obtained for $\Lambda = -m$ and its value is

$$R_{min} = 1$$

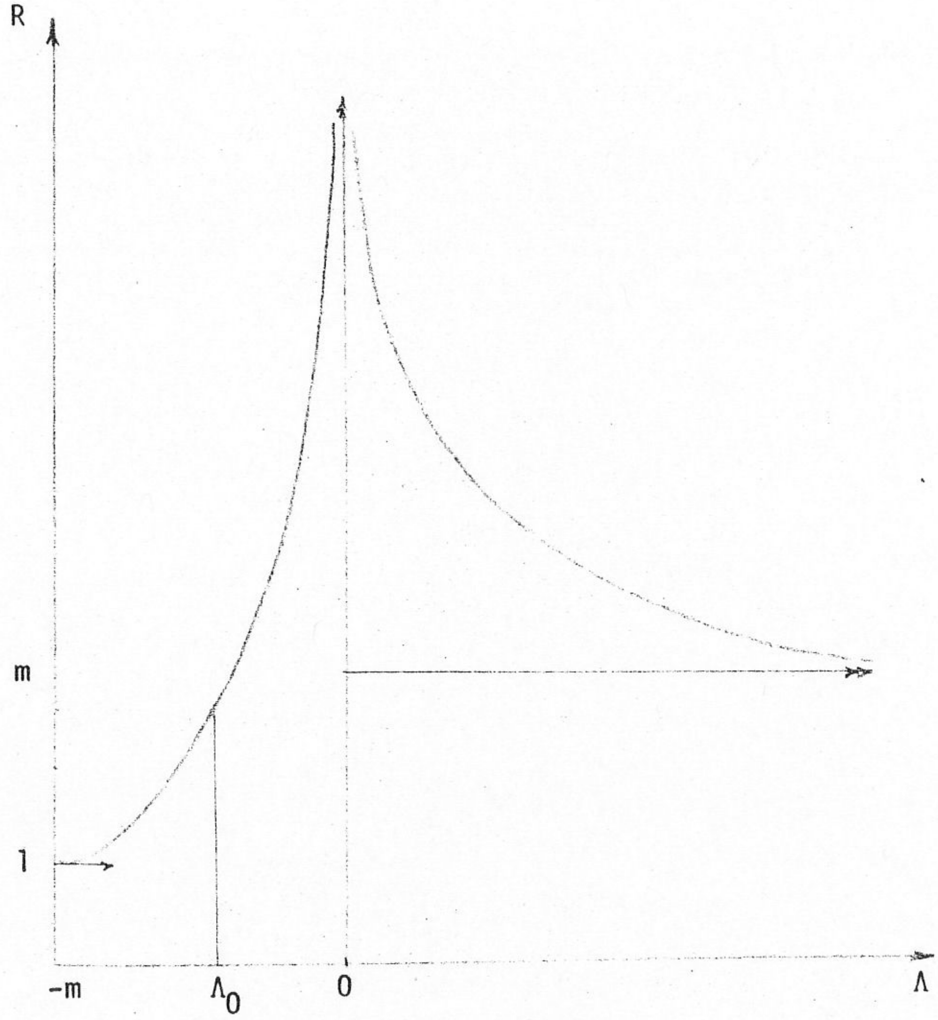

Figure 1. The function $R(\Lambda)$.

Since

$$2 + \frac{\sigma_e^2}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 \, f_1(\theta)d\theta + \sigma_e^2} < 3 \qquad (152)$$

in order to have the inequality (149) it is necessary and sufficient to have

$$\Lambda > \Lambda_0 > -m \qquad (153)$$

where Λ_0 is such that

$$2 + \frac{\sigma_e^2}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 \, f_1(\theta)d\theta + \sigma_e^2} = m\left(1 + 2\,\frac{m-1}{\Lambda_0} + m\,\frac{m-1}{\Lambda_0^2}\right) \qquad (154)$$

Now it is important to check that there exists a design f^* such that

$$\begin{cases} d_{f^*}(\theta) = m \dfrac{\sigma_y^2 + \sigma_x^2 \, |G(\theta)|^2}{\sigma_y^2 + \sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 \, f^*(\theta)d\theta} \\ \text{and} \\ d_{f^*}(\theta) > \Lambda_0 + m \end{cases} \qquad (155)$$

The first condition can be written as

$$\frac{d_f*(\theta) - m}{m} = \frac{\sigma_x^2 \left(|G(\theta)|^2 - \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta\right)}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta + \sigma_e^2} \tag{156}$$

the second condition being then

$$m + 2(m-1) \frac{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta + \sigma_e^2}{\sigma_x^2 \left(|G(\theta)|^2 - \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta\right)} + (m-1)\left[\frac{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta + \sigma_e^2}{\sigma_x^2 \left(|G(\theta)|^2 - \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta\right)}\right]^2 > 2 + \frac{\sigma_e^2}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta + \sigma_e^2} \tag{157}$$

or

$$1 + (m-1)\left[1 + \frac{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta + \sigma_e^2}{\sigma_x^2 \left(|G(\theta)|^2 - \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta\right)}\right]^2 > 2 + \frac{\sigma_e^2}{2\sigma_x^2 \int_0^{2\pi} |G(\theta)|^2 f^*(\theta)d\theta + \sigma_e^2} \tag{158}$$

which is automatically verified for $m \geq 3$.

It is now possible to summarize these results in the

Theorem VI. The three conditions are equivalent:

i) The design f^* is D-optimal for the C_3-constrained problem,

ii) f^* minimizes $\max_\theta \Delta_f^3(\theta)$,

iii) $\Delta_{f^*}^3(\theta) \leq 0$ for all $\theta\varepsilon[0,2\pi]$.

Proof. Because of (139), the proof is exactly as in Theorem III.

An iterative method of construction of C_3-constrained D-optimal designs proceeds exactly as in the preceding case. Its convergence--verified on numerical examples--has not been proved yet: since

$$(\sigma_x^2)_n^2 \int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta + \sigma_{x,n}^2 \sigma_e^2 = c_3 \tag{159}$$

$$\sigma_{x,n}^2 = \frac{-\sigma_e^2 + \sqrt{\sigma_e^4 + 4c_3 \int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta}}{2 \int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta} \tag{160}$$

and for

$$f_{n+1} = (1-\alpha)f_n + \alpha\, \delta_{\theta_n} \tag{161}$$

$$\frac{\sigma^2_{x,n+1}}{\sigma^2_{x,n}} = \frac{-\sigma_e^2 + \sqrt{\sigma_e^4 + 4c_3\left\{(1-\alpha)\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta + \alpha\, |G(\theta_n)|^2\right\}}}{-\sigma_e^2 + \sqrt{\sigma_e^4 + 4c_3 \int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta}}$$

$$\frac{\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta}{(1-\alpha)\int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta + \alpha\, |G(\theta_n)|^2} \tag{162}$$

And one has, after some computations:

$$\text{Det}\,(I_p(f_{n+1})) = R'(\alpha)\ \text{Det}\,(I_p(f_n)) \tag{163}$$

with

$$R'(\alpha) = R(\alpha)\left[\frac{\sqrt{2(\sigma_y^2-\sigma_e^2)^2 + 4\alpha c_3(|G(\theta_n)|^2 - \int_0^{2\pi} |G(\theta)|^2 f_n(\theta)d\theta)} - \sigma_e^2}{2(\sigma_y^2 - \sigma_e^2)}\right]^m \tag{164}$$

and

$$R(\alpha) = 1 + \frac{\alpha}{1-\alpha} d_{f_n}(\theta_n) \Bigg/ \left[1 + \frac{\alpha}{1-\alpha}\ \frac{|G(\theta_n)|^2}{\int_0^{2\pi} |G(\theta_n)|^2 f_n(\theta)d\theta}\right]^m \tag{165}$$

making an explicit determination of the value α_n such that $R'(\alpha_n)$ maximum extremely difficult[11].

The proof of the generalization of Kiefer's inequality presented in IV.1. is valid in the present situation, giving the

Lemma: Let f_x be in the neighborhood of an optimal design f_x^*, then

$$\frac{\text{Det}\,(I_p(f_x))}{\text{Det}\,(I_p(f_x^*))} \geq \exp\left\{- \max_\theta \Delta^3_{f_x}(\theta)\right\} \tag{166}$$

V. Existence of Stochastic and Mixed Solutions

The different versions of the equivalence theorem all express the fact that a necessary and sufficient condition for a design to be optimal is that

$$\Delta_{f^*}(\theta) \leq 0 \qquad \theta \varepsilon [0,2\pi] \tag{167}$$

on the other hand, one has

---

[11] One has the equivalent expression of $R'(\alpha)$:

$$R'(\alpha) = \left\{ \frac{1 + \sqrt{1 + 4K \int_0^{2\pi} |G(\theta)|^2 \, f_n(\theta) d\theta}}{1 + \sqrt{1 + 4K \int_0^{2\pi} |G(\theta)|^2 f_n(\theta) d\theta + 4\alpha K(|G(\theta_n)|^2 - \int_0^{2\pi} |G(\theta)|^2 f_n(\theta) d\theta)}} \right\}^{-m} \tag{165'}$$

with $K = \frac{C_3}{\sigma_e^2}$.

$$\int_0^{2\pi} \Delta_f^*(\theta)\, f^*(\theta) d\theta = 0 \tag{168}$$

implying that, at the points of the design $(f^*(\theta) > 0)$, one has

$$\Delta_f^*(\theta) = 0 \tag{169}$$

Now when one considers stochastic and mixed designs the continuous component of the spectral density is not a.e. equal to zero. For the sake of simplicity, assume that $f_x(\theta) > 0$ for all θ, this implies the very strong condition that

$$\Delta_f^*(\theta) = 0 \qquad \theta\varepsilon[0,2\pi] \tag{170}$$

Consider now the detailed form of $d_f^*(\theta)$:

$$d_f^*(\theta) = C \frac{1}{|\delta(\theta)|^4 \; f_e(\theta)} \operatorname{Tr} \left\{[I_p(f^*)]^{-1} \; [M(\theta)]\right\} \tag{171}$$

where

$$M(\theta) = (Q_{m;\mu,\nu}(\cos\theta, \sin\theta) \; ; \quad \mu,\nu = 1,\ldots,m) \tag{172}$$

where Q_m, $Q_{m;\mu,\nu}$ are polynomials of degree $m-1$ in $\cos\theta$, $\sin\theta$

$$d_f^*(\theta) = \frac{C\, Q_m(\cos\theta, \sin\theta)}{|\delta(\theta)|^4\, f_e(\theta)} \tag{173}$$

The existence of stochastic and mixed designs will now be discussed in the simplest case of the constraint C_1.

Using (173), $\Delta_f^1{}^*(\theta) = 0$ implies that:

$$C\, Q_m(\cos\theta, \sin\theta) = m\, |\delta(\theta)|^4\, f_e(\theta) \tag{174}$$

i.e., a necessary condition for the existence of a stochastic or mixed solution is that $|\delta(\theta)|^4\, f_e(\theta)$ is a "polynomial" in $\cos\theta$, $\sin\theta$ of degree at most $(m-1)$.

In the case of constraint C_2, (175) turns out to be:

$$C\, Q_{m-1}(\theta) = m\, C'|\omega(\theta)|^2\, |\delta(\theta)|^2\, f_e(\theta) \tag{175}$$

and for the constraint C_3

$$C\, Q_{m-1}(\theta) = m\, C''\left|\sigma_y^2\, |\delta(\theta)|^2 + \sigma_x^2\, |\omega(\theta)|^2\right| |\delta(\theta)|^2\, f_e(\theta) \tag{176}$$

[Note that, since the error process is real and usually is modelled as an ARMA process, $f_e(\theta)$ is the ratio of two polynomials in $\cos\theta$. In addition, $|\omega(\theta)|^2$, $|\delta(\theta)|^2$ are polynomials in $\cos\theta$: the problem of optimal design is the problem of finding a function f^* such that the coefficients of the matrix $[I_p(f^*)]^{-1}$ satisfy the identity relations implicit in (174, 175, or 176).]

The above results are summarized in the

Theorem VII. A necessary condition for the existence of a stochastic or mixed D-optimal solution is that the following quantities are polynomials of degree at most $(m-1)$ in $\cos\theta$, $\sin\theta$

i) For constraint C_1: $|\delta(\theta)|^4 f_e(\theta)$

ii) For constraint C_2: $|\omega(\theta)|^2 |\delta(\theta)|^2 f_e(\theta)$

iii) For constraint C_3: $[\sigma_y^2 |\delta(\theta)|^2 + \sigma_x^2 |\omega(\theta)|^2] |\delta(\theta)|^2 f_e(\theta)$

VI. Conclusion

Except for the proof of the convergence of the iterative method for the construction of C_3-constrained D-optimal designs this report gives a complete answer to the first step of what could be a theory of optimal designs for dynamic models: under realistic assumptions it is shown that a design f_x is equivalent to a real function of one variable Δ_{f_x}, which is in turn very helpful not only in characterizing optimal designs, but also in constructing them.

References

[1] C. L. Atwood. Sequences Converging to D-Optimal Designs of Experiments. (To be published).

[2] G. E. P. Box and W. Hunter. Sequential Design of Experiments for Non-Linear Models. IBM Scientific Computing Symposium in Statistics, (1965).

[3] G. E. P. Box and G. M. Jenkins. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, (1970).

[4] G. E. P. Box and G. Tiao. Bayesian Inference in Statistical Analysis. Reading: Addison-Wesley, (1972).

[5] Ph. Courrège, J. L. Philoche and P. Priouret. Regression Lineaire et Estimateurs des Moindres Carrés. Annales de l'Institut H. Poincarre B, VII, 4, (1971).

[6] P. G. Dhrymes, L. R. Klein and K. Steglitz. Estimation of Distributed Lags. International Economic Review, Vol II, 2, (June 1970).

[7] J. Durbin. Estimation of Parameters in Time Series Regression Models. J.R.S.S. B, 22, (1960).

[8] V. V. Fedorov. Theory of Optimal Experiments. (Translated and edited by E. M. Klimko and W. J. Studden).

[9] G. Fishman. Spectral Methods in Econometrics. Rand Corporation, Santa Monica, California, 1968.

[10] U. Grenander and M. Rosenblatt. Statistical Analysis of Stationary Time Series. Wiley and Sons, 1957.

[11] E. J. Hannan. Regression for Time Series. Proceedings of the Symposium on Time Series Analysis. M. Rosenblatt (ed.), Wiley and Sons, 1963.

[12] G. M. Jenkins and D. Watts. Spectral Analysis and Its Applications. Holden-Day, San Francisco, (1969).

[13] S. Karlin and W. J. Studden. Tchebycheff Systems: With Applications to Analysis and Statistics. J. Wiley and Sons, New York, (1966).

[14] J. Kiefer. Optimal Experimental Designs V, With Applications to Systematic and Rotatable Designs. Fourth Berkeley Symposium. J. Neyman Ed., (1961).

[15] W. Kruskal. When are Gauss-Markov and Least Squares Estimators Identical? A coordinate free approach. A.M.S., 39, (1968).

[16] J. MacGregor. Topics in Control II: Constrained Feed-Back Control of Linear Dynamic-Stochastic Systems. Technical Report No. 307, University of Wisconsin, Department of Statistics, (1972).

[17] G. M. Minnich. Some Considerations in the Choice of Experimental Designs for Dynamic Models. Ph.D. Dissertation. Department of Statistics, University of Wisconsin, (1972).

[18] E. Parzen. An Approach to Time Series Analysis. A.M.S., Vol 32, 4, (Dec. 1961).

[19] W. Rudin. Real and Complex Analysis. McGraw-Hill, (1966).

[20] R. St. John. D-Optimality: A Review. Master's Thesis. Department of Statistics, University of Wisconsin, Madison, Wisconsin, (1971).

[21] B. L. Viort. On the Dispersion Matrix of the Maximum Likelihood Estimators of the Parameters of a Dynamic Model. Technical Report No. 313, University of Wisconsin, Department of Statistics, (1972).

[22] G. Wahba. Estimation of the Coefficients in a Multidimensional Distributed Lag Model. Econometrica, Vol. 37, 3, (July 1969).

[23] G. S. Watson. Linear Least Squares Regression. A.M.S., 38, (1967).

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

University of Wisconsin
Department of Statistics
Madison, Wisconsin

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

D-Optimal Designs for Dynamic Models. Part I: Theory

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Bernard Viort

6. REPORT DATE

October 1972

7a. TOTAL NO. OF PAGES

53

7b. NO. OF REFS

23

8a. CONTRACT OR GRANT NO.

AFOSR-72-2363

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

Technical Report Number 314

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Air Force Office of Scientific Research
1400 Wilson Boulevard
Arlington, Virginia

13. ABSTRACT

The problem of design of experiments for dynamic models is very difficult to solve in the time-domain. The rotation to the frequency domain reduces the problem to a well-known set-up and makes available many results and methods of the now classical theory of design of experiments.

Theoretical results as well as numerical algorithms are given for different types of constraints on the input and/or output variances.

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
D-optimality						
Dynamic models						
Design of experiment						
Constraint						

Unclassified
Security Classification