CONTROL OF CRITICAL SYSTEMS

by

Theodoros Lokkas

A dissertation submitted to

THE VICTORIA UNIVERSITY OF MANCHESTER

in partial fulfilment of the requirements

for the degree of

MASTER OF SCIENCE

Supervisor: Dr. V. Zakian

Control Systems Centre
The University of Manchester
Institute of Science & Technology

NOVEMBER 1986

Dedicated to

My Wife MARY

and

My Daughter EVITA

SUMMARY

There are two main chapters in this dissertation and each of them has its own objective. The first objective is to present some of the theoretical aspects of Zakian's design framework and its application to the design of critical systems. The second one is to show how these aspects can be applied to the design of a critical system.

Analytically, the 1st chapter includes a brief outline of Zakian's approach to control system design. It contains a formulation of the theory and explanations of the input space and the performance criterion for both rational and non-rational feedback systems. There is also a description of the CRITERIA software package, which is a computer-aided control system design package incorporated in Zakian's framework, with a summary of the process of design using this package for both rational and non-rational systems. There then follows a discussion on the significance of the various types of critical systems and their relationship to Zakian's design criterion. Finally, there are descriptions on how the input space for random inputs can be estimated from measurements as well as from stochastic properties for stationary processes. In the latter case there are some explanations of the stochastic process theory necessary for the formulation of the stochastic properties.

The 2nd chapter describes an actual critical system - highway automatic car-following control - and the design of this system using Zakian's criteria. There is first a general outline of ground transportation systems and then follows the special description of the car-following system with its requirements. The model analysis for one of the numerous controllers - a linear velocity controller - is derived, as well as the input space for the system

is determined and the criteria specifications for a variety of controllers are exposed, which would be used in the system. Finally, three kinds of compensators are designed for the system using CRITERIA, and a comparison is made in the system design between Zakian's method and a traditional one in the case study.

CONTENTS

INTRODUCTION

Critical systems are the systems wherein the system error must be contained within certain bounds, because otherwise the consequences could be dramatic or unpleasant.

There are different types of critical systems depending mainly upon the degree of criticality. For example, a very critical system might be one where the results would be catastrophic if the system error exceeded a permissible value - the critical bound.

Such systems like the landing control of an aeroplane, the velocity control of a train or car, especially in the curves of the rails or roadways respectively, and generally every control system wherein people co-exist, are indeed very critical ones because losses of lives might occur if the error bounds were exceeded.

Anyway, regardless of less or more critical systems, the design framework based on Zakian's criteria is well suited to designing critical systems, because it provides both an explicit definition of control and takes into account the system input.

In all conventional methods an input such as a step function is assumed, which does not exist in nature because systems actually are subjected to inputs which vary enormously, such as random disturbances. Zakian's design criteria considers a class of inputs - elements of a known input space - with a simple property - the bound on the input derivative - characterized as D. This input space is incorporated in the design process so that the effect of the input on the system performance is known throughout the process. The problem is to determine the input space.

This dissertation describes the process of the design of critical systems with single input/single output using Zakian's design criteria. Especially, it shows how the input space required can be estimated both from measurements of stochastic inputs and from the properties of stochastic inputs, for stationary processes, as well as how the design criterion is determined and how the system is designed using a computer-aided design package. One complete design of an actual critical system is described using Zakian's framework.

Chapter 1

ZAKIAN'S THEORY AND ITS IMPLEMENTATION

1.1 Formulation of the Theory

The mathematical formulation of Zakian's theory [42] is based on the fact that the control system is represented by a pair (S,F) where F is the input function space and S is the input-output rule

$$S : \quad f \rightarrow e(f,p) \tag{1.1}$$

This rule maps the input space F into an output function space, where

$f \in F$ is the input function

$e(f,p)$ is the output function and

$p = (p_1, p_2, \ldots, p_n) \in R^n$ is the design parameter

(e.g. gain of an amplifier, damping coefficient, etc.)

The input f is known only to the extent that it belongs to the known function space F, which is defined as the class of all functions $f : R \rightarrow R$ such that

$$f(t) = 0 \quad \text{for all} \quad t \leq 0 \tag{1.2}$$

and there is a positive finite number D such that

$$D = \sup\{|f^{(1)}(t)| : t \geq 0 \, , \quad f \in F\} \tag{1.3}$$

i.e. D is the greatest absolute rate of change of the input functions as time t ranges over the real numbers and the input f ranges over the space F. Here $f^{(1)}(t)$ represents the derivative of $f(t)$, i.e.

$$f^{(1)}(t) = \frac{df(t)}{dt} \tag{1.4}$$

which is piecewise continuous. The piecewise continuity of $f^{(1)}(t)$ is necessary so that it can be integrable over its interval.

From (1.3) it is clear that the input space F is characterised by its finite number D. The output is the error function

$$e(f,p) : t \rightarrow e(t,f,p) \tag{1.5}$$

which maps the time interval R into the real line R and depends on a parameter $p \in R^n$.

It is assumed that the input-output function

$$f \rightarrow e(f,p)$$

which maps the input space F into an output function-space, is causal, linear and time-invariant.

Now, let $h : R \rightarrow R$ denote the unit step function

$$h(t) = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } t \geq 0 \end{cases} \tag{1.6}$$

and $e(h)$ denotes the unit step response of the system $(S,\overline{F})$ (i.e. the error function when $\overline{f}$ is the unit step function h). At time t the value of the step response is $e(h,t)$ which is assumed to be piecewise continuous.

The error function (1.5) of the system is defined in a general form by the convolution integral

$$e(t,f,p) = \int_0^t e(t-\lambda,h,p)f^{(1)}(\lambda)d\lambda \tag{1.7}$$

For example, if we consider a feedback system of standard form, characterised by the block diagram in Figure 1.1, the error function (1.5) is defined, in terms of its Laplace transform, by

$$L[e(t,f,p)] = \underline{E}(s,f,p) = \frac{F(s)}{1+G(s)K(s,p)} \tag{1.8}$$

where s denotes the Laplace transform variable,

$F(s)$ the transform of the input function $f(t)$,

$G(s)$ the transfer function of the plant, and

$K(s,p)$ the transfer function of the controller, which depends on the parameter $p \in R^n$.

If the output $e(f,p)$ is taken as the error of the system, a natural measure of the lack of performance of the system $\phi(p)$ can be defined as

$$\phi(p) = \sup\{|e(t,f,p)| : t \geq 0 , \quad f \in F\} \tag{1.9}$$

In fact $\phi(p)$, which is called performance functional, is the worst error as time t ranges over the half line $[0,\infty)$ and the input f ranges over the input space F.

1.2 Design and Stability of Control Systems

The performance functional

$$\phi : p \to \phi(p) \geq 0 \tag{1.10}$$

which maps $p \in R^n$ into the extended half-line $[0,\infty]$ is such that for some $p \in R^n$, $\phi(p)$ can be infinite.

It is obvious that given a finite positive number ε (specified by the designer), a criterion of design requires that the controller parameter p has to be chosen so that the inequality

$$\phi(p) \leq \varepsilon \tag{1.11}$$

is satisfied. The number ε represents the worst performance that can be tolerated. In this case we say that the value of p characterises an acceptable design. The design criterion (1.11) provides an explicit statement of the extent of control required over the error $e(f,p)$.

A solution of (1.11) can be achieved by numerical methods and involves two phases of computations. The first phase consists of a generation of a finite sequence of points $p \in R^n$, which terminates with a point p such that

$$\phi(p) < \infty \tag{1.12}$$

In this case, since $\phi(p)$ is a finite number, the system can be characterised as stable and the set of $p \in R^n$ for which $\phi(p) < \infty$ constitutes the stability region.

The second phase comprises the generation of a finite sequence which starts inside the stability region and terminates at inequality (1.11). The set of $p \in R^n$, for which $\phi(p) \leq \varepsilon$, constitutes the admissible region.

However, the second phase of computation is a problem of economy. The evaluation of the functional $\phi(p)$ at each point p is costly if it is carried out from the expression (1.9) mainly because of the supremum operation over the input space F .

Unfortunately the stability problem cannot always be solved by numerical methods because:

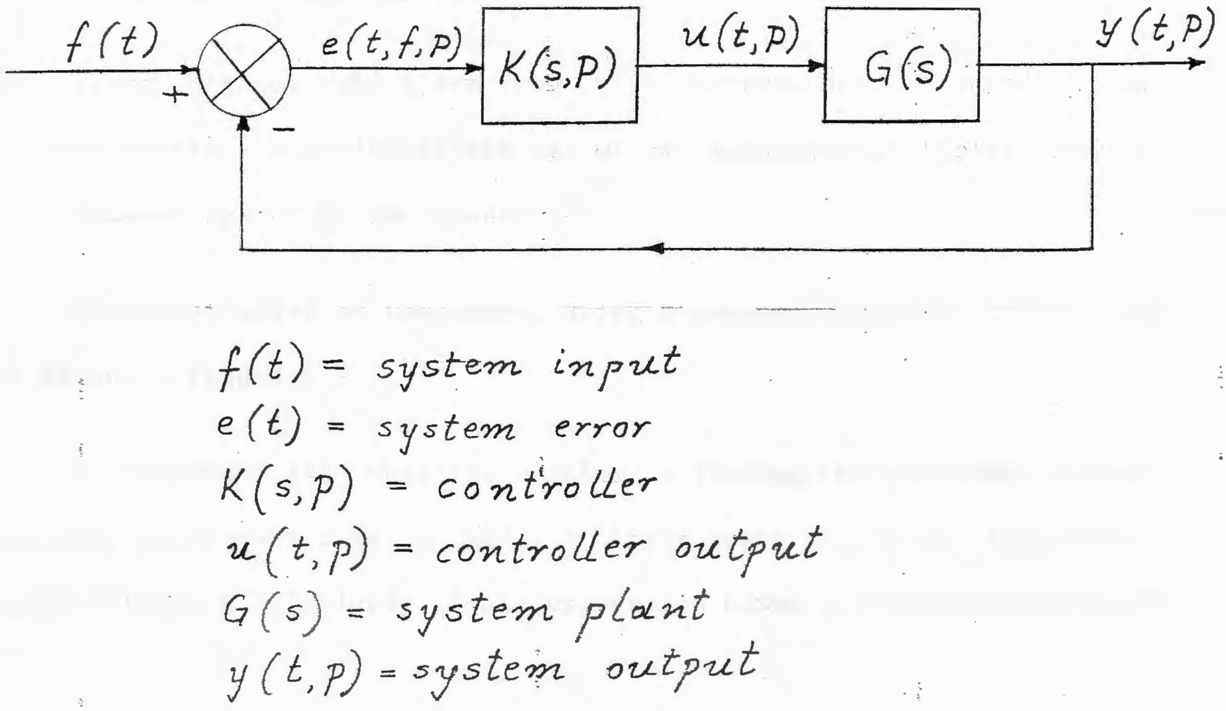


Figure 1.1 Standard unity feedback system.

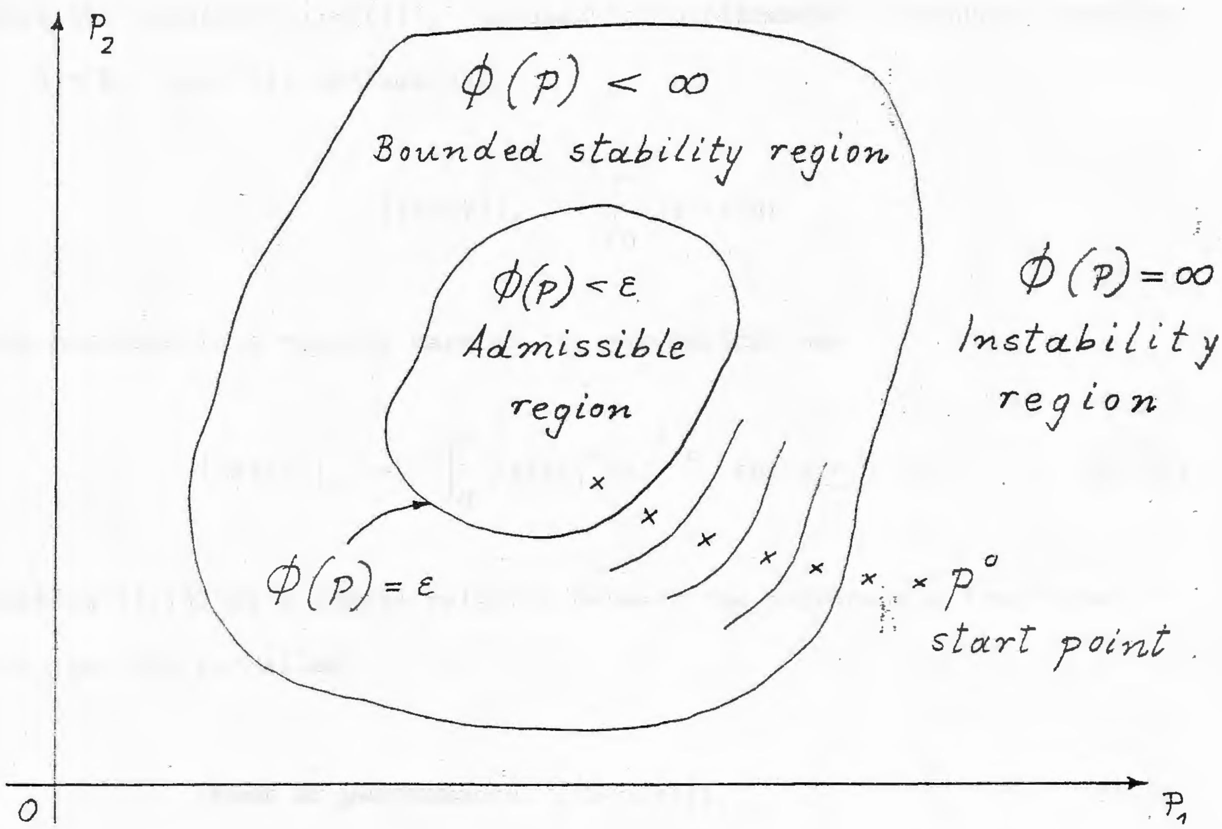


Figure 1.2 Regions for a control parameter $p = (p_1, p_2) \in R^2$.

(a) descent methods cannot be used (a gradient or similar property of ϕ cannot be defined outside the region defined by (1.12), and

(b) search methods (which are less efficient than descent methods) can fail to locate a solution if the set of all solutions of (1.12) forms a bounded region in the space R^n .

An illustration of the above, using a control parameter $p = (p_1, p_2) \in R^n$ is shown in Figure 1.2 .

It is proved [42] that the problem of finding the supremum of the absolute error over time t and over input space F, $\phi(p)$ is solved explicitly by a relatively simple expression known as Fundamental Theorem,

$$\phi(p) = ||e(h,p)||_1 \; D \tag{1.13}$$

where the notation $||x(t)||_1$ is used for a piecewise continuous function $x : R \to R$, which is defined as

$$||x(t)||_1 = \int_0^\infty |x(t)|dt \tag{1.14}$$

This notation is a special case of the generalized one

$$||x(t)||_n = \{\int_0^\infty |x(t)|^n dt\}^{1/n} \quad \text{for } 1 \leq n < \infty \tag{1.15}$$

Equation (1.15) is a simple relation between the performance functional (p) and the so-called

$$\text{index of performance} \quad ||e(h,p)||_1 \tag{1.16}$$

which is referred to as the Integral of Absolute Error (IAE).

1.3 Stability of Rational Feedback Systems

An equivalent condition for the satisfaction of inequality $\phi(p)$ is provided by the following theorem [41] which is valid for the stability of rational systems (known as Stability Theorem) :

$\phi(p)$ is finite iff:

(a) $E(s,\delta,p)$ has integral action, and

(b) $\text{Re}\,\lambda(p) < 0$ for all $\lambda(p) \in \Lambda(p)$,

where

$$E(s,\delta,p) = \frac{1}{1+G(s)K(s,p)} \tag{1.17}$$

is the error transfer function which is supposed to be a rational function of s, δ is the Dirac delta function, $\Lambda(p)$ denotes the set of the poles $\lambda(p)$ of $E(s,\delta,p)$, and $\text{Re}\,\lambda(p)$ denotes the real part of the pole $\lambda(p)$.

Notes:

1. Integral action of the rational function $E(s,\delta,p)$ means that it has a zero at $s = 0$, i.e.

$$E(0,\delta,p) = 0 \tag{1.18}$$

For a feedback system of the form (1.17), an obvious necessary and sufficient condition for integral action is

$$G(0)K(0,p) = \infty \tag{1.19}$$

2. Since $E(s,\delta,p)$ is a rational function of s, then $E(\infty,\delta,p)$ is equal to zero or to a finite real number, whether the degree of the numerator is less or equal to the degree of the denominator respectively. Hence $e(1,p)$ is continuous and $||e(1,p)||_1$ is well-defined.

The importance of this theorem is that $\mathrm{Re}\,\lambda(p) < 0$ is always finite, whereas $\phi(p)$ may be infinite for some points $p \in R^n$. Therefore condition (b) of the theorem can be solved by gradient methods, while inequality $\phi(p) < \infty$ cannot frequently be solved by numerical methods.

1.4 Stability of Non-Rational Feedback Systems

We consider the unity feedback system of Figure 1.1 where the controller transfer function $K(s,p)$ is a proper rational function of s, but the plant transfer function $G(s)$ is not. Then the transfer function (1.8) will not be a rational function of s, which means that the previously discussed stability theorem will not hold. So we must find a means which can satisfy the stability criterion $\phi(p) < \infty$ for non-rational systems.

Let $G^*(s)$ denote a strictly proper rational function (i.e. $G^*(\infty) = 0$) called nominal plant transfer function or rational approximant of $G(s)$. If $G(s)$ is replaced with $G^*(s)$, then the resulting system is called a nominal feedback system. This system has a nominal error $e^*(f,p)$ which is related to the input f by the transfer function

$$E^*(s,\delta,p) = \frac{1}{1+G^*(s)K(s,p)} \qquad (1.20)$$

whose finite poles $\lambda^*(p)$ constitute the set $\Lambda^*(p)$. The performance of a nominal system is measured by

$$\phi^*(p) = \sup\{|e^*(t,f,p)| : t \geq 0,\ f \in F\} \qquad (1.21)$$

Now define

$$\mu(p) = ||w(p)||_1 \qquad (1.22)$$

where

$$w(p) : t \to w(t,p) \text{ has Laplace transform}$$

$$W(s,p) = X(s,p)Z(s) \qquad (1.23)$$

and $$Z(s) = G(s) - G^*(s) \tag{1.24}$$

$$X(s,p) = \frac{K(s,p)}{1+G^*(s)K(s,p)} \tag{1.25}$$

Let $\lambda_u^*(p)$ denote the finite poles of $X(s,p)$ and let these poles constitute the set $\Lambda_u^*(p)$.

It has been shown by Zakian [42], as in the case of rational systems, that the following three conditions are sufficient in a non-rational system for $\phi(p) < \infty$,

(a) $E^*(0,\delta,p) = 0$

(b) $\text{Re}\ \lambda^*(p) < 0$ for all $\lambda^*(p) \in \Lambda^*(p)$

(c) $\mu(p) < 1$.

The number $\mu(p)$ is a measure of the difference between the nominal system and the original system.

The function $\mu : p \rightarrow \mu(p)$ maps R^n into the extended half-line $[0,\infty]$. Therefore there may be points $p \in R^n$ such that $\mu(p)$ is infinite. So the first step of the computational procedure to satisfy condition (c) is to generate a sequence of $p \in R^n$ which satisfy the inequality

$$\mu(p) < \infty \tag{1.26}$$

However, in the same way as the inequality $\phi(p) < \infty$, condition (1.26) cannot be satisfied by numerical methods given only the function $\mu(p)$. Thus, we must find new conditions which can be satisfied by numerical methods to replace the inequality (1.26).

In the same way Zakian has proved [42] that the following two conditions are sufficient for $\mu(p) < \infty$,

(d) $||z||_1 < \infty$

(e) Re $\lambda_u^*(p) < 0$ for all $\lambda_u^*(p) \in \Lambda_u^*(p)$.

It is usually found in practice that $G^*(s)$ can be chosen so as to make $||z||_1$ arbitrarily close to zero [2] . This implies that provided condition (e) is satisfied, $\mu(p)$ can be made arbitrarily close to zero.

Since the relationship between $\phi(p)$, $\phi^*(p)$ and $\mu(p)$ is [42]:

$$\frac{\phi^*(p)}{1+\mu(p)} \leq \phi(p) \leq \frac{\phi^*(p)}{1-\mu(p)} \tag{1.27}$$

this means that $\phi^*(p)$ can be made arbitrarily close to $\phi(p)$.

Note that $\Lambda^*(p) \subseteq \Lambda_u^*(p)$. Because, if for example P_i, i = 1,2,3,4,5,6 denote polynomials such that

$$K(s,p) = \frac{P_1P_2}{P_3P_4}, \quad G^*(s) = \frac{P_5P_4}{P_6P_2}, \tag{1.28}$$

then $\Lambda^*(p)$ is the set of all zeros of $P_3P_6+P_1P_5$, while $\Lambda_u^*(p)$ is the set of all zeros of $P_4(P_3P_6+P_1P_5)$ as can be easily derived from (1.20) and (1.25) respectively.

1.5 Design of Rational and Non-Rational Systems

Using the above theory Zakian's group has developed an interactive computer-aided control system design package under the name CRITERIA [3] at the Control Systems Centre, UMIST, Manchester. This package is located on the PRIME 9950 computer at the UMIST Control Systems Centre.

CRITERIA embodies a number of simple but powerful algorithms which are used to solve the inequalities arising from Zakian's criteria. It has been specifically designed for easy and efficient user-defined upgradability and so forms a modern and comprehensive design tool, which can be used to solve a wide range of control design problems.

The package can be used for design of both rational and non-rational single-variable unity feedback systems.

So, the procedure for the design of single-input/single-output *rational* systems is as follows:

1. Define the open-loop transfer function written as quotient with numerator and denominator polynomials.

2. Determine the bound D on the input space F as defined in equation (1.3).

3. Specify the form of the compensator $K(s,p)$, then define the initial parameters of the compensator and specify any constraints on the values of the compensator.

4. Specify the finite positive number ε which is the greatest error tolerance and enter this upper bound ε of the performance functional $\phi(p)$ defined by (1.9). A bound is also required for the controller output $u(t)$.

5. By the method of inequalities calculate a point $p \in R^n$ that satisfies the condition

$$a(p) < 0 \tag{1.29}$$

$$\text{where} \quad a(p) = \max\{\text{Re}\,\lambda(p) : \lambda(p) \in \Lambda(p)\} \tag{1.30}$$

A number of iterations of an algorithm are performed varying the compensator parameters until the system becomes stable and $\phi(p) < \infty$. If the inequality $\phi(p) \leq \varepsilon$ is not already satisfied, a new number of iterations of an algorithm can be performed until the design criteria are satisfied.

At last we obtain a list of the parameters $p \in R^n$, the error functional $\phi(p)$, the control functional $\phi_u(p)$ given by

$$\phi_u(p) = \sup\{|u(t,f,p)| : t \geq 0, \ f \in F\} \qquad (1.31)$$

and the abscissa of stability $a(p)$ given by (1.30).

We can also obtain plots of the system unit step response, the controller unit step response and the absolute error unit step response.

The designer may change or improve the formulation of the problem by increasing the complexity of the controller, providing thus a more expensive implementation, or loosening the bound ε, giving a less stringent specification. He may also decrease the controller complexity, giving a cheaper implementation, or tighten the bound ε, obtaining a higher quality system if $\phi(p)$ is well within the specified bound ε.

The procedure for the design of single-input/single-output *non-rational* systems is slightly modified as follows:

1. Define the non-rational open-loop transfer function $G(s)$ of the system, written as quotient with numerator and denominator polynomials times a non-rational component.

2. Determine a rational approximant $G^*(s)$ of the original transfer function $G(s)$.

3. Determine the bound $\bar{D}$ on the input space F as defined in equation (1.3).

4. Specify the form of the compensator $K(s,p)$. Then define the initial parameters of the compensator and specify any constraints on the values of the compensator.

5. Specify the finite positive number ε and enter this upper bound ε of the performance functional $\phi(p)$ as well as a bound for the approximation index $\mu(p)$ defined by (1.22). It is required also a bound for the control output $u(t)$.

6. By the method of inequalities locate a point $p \in R^n$ that satisfies condition (e) of non-rational systems to ensure that $\mu(p) < \infty$ and then reduce $\mu(p)$ until $\phi(p) < \infty$. Finally, reduce $\phi(p)$ until the design criterion $\phi(p) \leq \varepsilon$ is satisfied.

At the end we obtain a list of the parameters $p \in R^n$, the error functional $\phi(p)$, the control functional $\phi_u(p)$, the abscissa of stability $s(p)$ and the approximation index $\mu(p)$, as well as the integral absolute error of the approximant system.

We also can obtain plots of the system unit step response, the controller unit step response and the absolute error unit step response.

1.6 Zakian's Theory for Critical Systems

A critical system, as mentioned in the introduction, is one where the system error must not exceed a maximum allowable value, for otherwise it could have unpleasant or even disastrous results. From the definition itself, it becomes quite clear that the use of the design criterion specified by $\phi(p) \leq \varepsilon$, where ε is the greatest error that we can tolerate, makes Zakian's control theory ideal for the design of critical systems.

The characterization of a system as a critical one calls for some explanation. In fact, there are some factors which play a significant role for the degree of "criticality" of a system, such as personal or non-personal security, health, comfort, capacity, economy, inflation, etc. Thus, according to the degree of "severity" (i.e. the various factors included in the system), there are very critical systems and less critical ones.

A very critical system may be one wherein the consequences of the error exceeding the permissible bound are very tragic. In an aircraft, for example, the systems for landing or taking off are very critical, because accidents or even loss of lives may occur if the error bounds are exceeded. The same case appears in a spacecraft when it enters from the space into the atmosphere. Great (relatively) values of entering-angle could cause a very dangerous increase of the spacecraft's temperature because of friction effects; small ones could force it going away from atmosphere. Generally, every moving fast transport vehicle (car, train, etc.), which carries people, can (as a whole) be considered as a very critical system.

Other systems may not have disastrous but still highly undesirable results. Such systems are a satellite tracking antenna, where communication errors or even loss of tracking will result if the permissible error is exceeded. An analogous very simple system is our effort to keep having a straight rod with our finger perpendicular.

There are systems where the consequences of the error bound being exceeded for a short time or by a small amount are not too catastrophic, provided that the error remains less than ε for most of the time. Such systems are the pollution control for a river, the pollution control for the air, etc.

It may be that the actual permissible error is known exactly. For example, the temperature maximum error in hospitals, schools or in public buildings, is a statutory requirement defined by law, or the temperature maximum error in greenhouses defined by the temperature requirements of the plants. Another example of a more critical system is a highway automation system where the spacing between two vehicles moving in the same direction and lane must be kept as small as possible so that on the one hand to avoid a collision, and on the other hand to get as great as possible capacity. This system is described and designed in the next chapter.

However, there are cases where the bound is roughly known and no exact number can be given to ε so that the maximum permissible error may have some obstruseness about it, depending to some extent on quality requirements, like a paper machine .

In any of the different cases of critical systems, the problem of determining the bound ε of the design criterion is not always straightforward, in most cases often depending mainly upon safety or quality requirements.

1.7 Determination of the Input Space F

A. From Measurements

The design criteria outlined above for the rational and non-rational systems are significantly different from other classical approaches. For example, in the classical control theory the word "system" has the meaning of the input-output rule (e.g. a transfer function). In Zakian's theory this word has a larger meaning, which incorporates a carefully chosen input space.

Thus, for an engineer applying Zakian's criteria to the design of a control system, the determination of the number D, which represents the supremum of the absolute rate of change of the input, is one of the most important tasks, and the significance of D increases with the degree of criticality of the system.

So in most aeronautical systems, where the exceedance of the design criterion could be fatal and result in many losses of lives, D is very important. In any case, irrespective of whether the systems are very critical or less critical, the correct calculation of D will ensure or contribute to a better design.

The best means to determine the bound D on the input space is from measurements of the maximum rate of change of the input.

For the determination of an input $f_1(t)$ which belongs to an input space F, a number of samples of the contents of $f_1(t)$ are taken at different time intervals, and these samples form an analytical expression for the input $f_1(t)$. The maximum of the derivative of $f_1(t)$ constitutes an element of the set needed for the determination of D. The same procedure is repeated for another input $f_2(t)$ and so on. The maximum element of all the samples of $f_1, f_2, \ldots, f_n$ thus obtained characterises the number D.

But even taking D as the largest measured extreme for all the samples $f_1, f_2, \ldots, f_n$ over time t, it is not always adequate for critical systems because, no matter how large the sample $f_1, f_2, \ldots, f_n$ is, there is always the possibility that the greatest rate of change of $f(t)$ of all these samples could be exceeded simply because the sample size is finite.

However, there are methods for making a good estimate of measured extreme data, as well as for determining the probability that a certain value will be exceeded within a certain time. These methods are the subject of the statistical theory of extremes and can be very useful for the forecasting of natural process extremes, such as highest wind speeds, highest sea-waves, highest temperatures and so on. Thus, this theory can be used with very good results for the critical systems to obtain a value of D with a known amount of confidence.

The main problem with the statistics of extremes approach is that the data required for most disturbances to estimate D are rare. So, [19], at least 20 different sample periods are required and a value for the maximum observed random variable for each sample period. The period used must be one such that the overall conditions must be similar in each period.

Thus, for meteorological random variables, such as wind speed, a suitable period would be years or days in a particular month. Unsuitable periods would be hours or months. This because the conditions under which the maximum wind speed value is measured must have remained constant for the duration of the recordings, and the daily conditions in a particular month are likely to be similar. However, there is a great variation in the weather over different months.

So, records need to have been kept for 20 years if they are to be of use for a control engineer for designing very critical systems subjected to climate disturbances.

Of course, if the system is not so critical, such as the air pollution, the data need not have been kept for so long a time, since the value of D is not so important.

B. From Power-Density Spectrum for Stationary Processes

As has been mentioned above, the data required to estimate D by the statistical theory of extremes is not always available. However, it is sometimes possible to estimate D from some known properties, such as the power-density spectrum or spectral density function of the stochastic input.

Feedback systems are often subject to stochastic disturbances; if a stochastic input f(t) to the system is assumed to be stationary and Gaussian and an expression for the power-density spectrum is available, then it is possible to determine a suitable value for the bound D on the input space F.

It must be pointed out, however, that making the assumption that the random disturbance is a Gaussian stationary (defined in this paragraph) the originally defined input space F according to Zakian's theory is obviously not valid any more, or better, it becomes a subspace F_1 in which the calculation of D will be an approximation of the real one. However, we do make this assumption just because it is easier to have an approximate calculation of D than to wait for a long time to get some data from the statistics of extremes for disturbances like wind, rain, sea-waves, etc.

Such assumptions could be made only in the case where the control system under consideration was not a very critical one, such as the system discussed in Chapter 2; in the opposite case it would require a very exact calculation of D.

But first let us briefly have some definitions and explanations of the stochastic process theory for further study.

Let S be the sample space of some experiment. We assign a specific number to each outcome, for example the sum of the points on a pair of dice, etc. So we have a function on the sample space S which is called a random variable or stochastic variable (more precisely random function) usually denoted by a capital letter X or Y.

A random variable (r. v.) which takes on a finite or countably infinite number of values is called discrete r.v., while one which takes on a non-countably infinite number of values is called a continuous r.v.

Let the image set of a discrete r.v. X is $X(S) = \{x_1, x_2, \ldots, x_n\}$, and suppose that the x_i's are assumed with probabilities given by

$$f(x_i) = P(X = x_i) \tag{1.32}$$

This function f is called a probability function of X (or distribution of X) which satisfies the conditions:

$$\text{(i)} \quad f(x_i) \geq 0 \; . \quad \text{(ii)} \quad \sum_{i=1}^{n} f(x_i) = 1 \tag{1.33}$$

The (cumulative) distribution function for a random variable X is defined by:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) \tag{1.34}$$

If X is a continuous r.v. then (1.34) becomes:

$$F(x) = \int_{-\infty}^{x} f(t)dt \tag{1.35}$$

from which $$f(x) = \frac{dF(x)}{dx} \tag{1.36}$$

The probability that X lies between two values x_1 and x_2 is:

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx = F(x_2) - F(x_1) \tag{1.37}$$

Now, if X is a r.v. with the probability function (1.32) the mean or expectation, or expected value of X, denoted by $E(X)$ or m_X or simply m is defined by

$$m = E(X) = \sum_{i=1}^{n} x_i f(x_i) \tag{1.38}$$

for a continuous r. v.

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)dx \tag{1.39}$$

The mean of a random variable X measures the "average" value of X in some sense. The variance of X, denoted by $var(X)$ or σ_X^2 or simply σ^2 is defined by

$$var(X) = \sum_{i=1}^{n} (x_i - m)^2 f(x_i) = E[(X-m)^2] \tag{1.40}$$

and measures the "spread" or "dispersion" of X. Equation (1.40) can also be written as:

$$var(X) = E(X^2) - m^2 \tag{1.41}$$

or in continuous form:

$$var(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - m^2 \tag{1.42}$$

The standard deviation of X, denoted by σ_X or simply σ is the non-negative square root of var(X), i.e.

$$\sigma_X = \sqrt{\mathrm{var}(X)} \tag{1.43}$$

If X, Y are two r.v. we define the joint probability function of X and Y by

$$f(x,y) = P(X = x, Y = y) \tag{1.44}$$

where $f(x,y) \geq 0$ for every x, y and

$$\sum_x \sum_y f(x,y) = 1 \quad \text{(for discrete r. v.'s)} \tag{1.45}$$

or $$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1 \quad \text{(for continuous r. v.'s)} \tag{1.46}$$

The covariance of X and Y denoted by r(X,Y) is defined by

$$r(X,Y) = E[(X-m_X)(Y-m_Y)] \tag{1.47}$$

with the property

$$r(X,Y) = E(XY) - m_X m_Y \tag{1.48}$$

and the correlation of X and Y denoted by $\rho(X,Y)$ is defined by

$$\rho(X,Y) = \frac{r(X,Y)}{\sigma_X \sigma_Y} \tag{1.49}$$

Now, a stochastic process can be defined as a set of random variables $\{X(t) : t \in T\}$. The variables $X(t)$ are indexed with the parameter t, which means a specific time. For continuous time processes the time t belongs to the half-line $[0,\infty)$. The value of a stochastic process cannot be predicted exactly at any particular future time. However, the probability of the stochastic process can be predicted, lying within a particular range of values.

In a stochastic process the covariance is a more general case of the variance, defined as

$$r_X(t_1,t_2) = E\{[X(t_1)-m(t_1)][X(t_2-m(t_2)]\} \tag{1.50}$$

If the distribution of the random variables $X(t_1),X(t_2),\ldots,X(t_n)$ is identical to the distribution of $X(t_1+\tau),X(t_2+\tau),\ldots,X(t_n+\tau)$ for every τ such that $t_i \in T$ and $(t_i+\tau) \in T$, $i = 1,2,\ldots,n$, then the stochastic process $\{X(t), t \in T\}$ is said to be stationary. This means that the characteristics of the process remain constant for all time, if it is a stationary process. If only the mean and covariance of the distributions are constant, the process is said to be weakly stationary.

One of the most important continuous probability distributions is the normal or Gaussian distribution with a probability density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \tag{1.51}$$

where m is the mean and σ^2 is the variance of the distribution. Figures 1.3 and 1.4 show $f(x)$ with constant m and σ respectively.

For $m = 0$, $\sigma^2 = 1$, equation (1.51) becomes:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{1.52}$$

with a graph shown in Figure 1.5 .

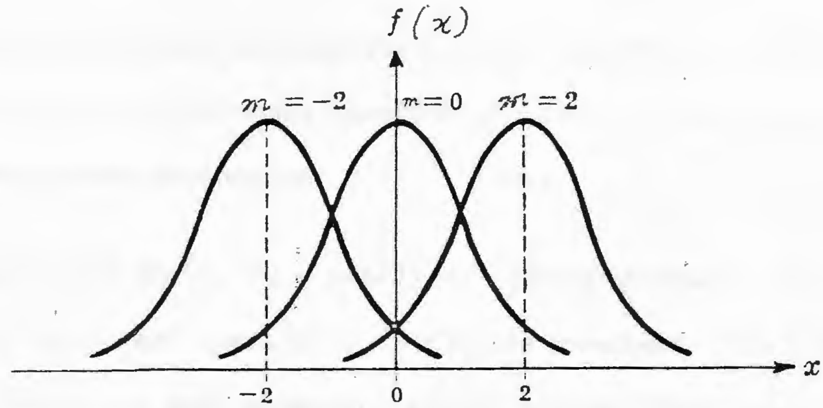


Figure 1.3 Normal distributions with σ fixed (σ = 1)

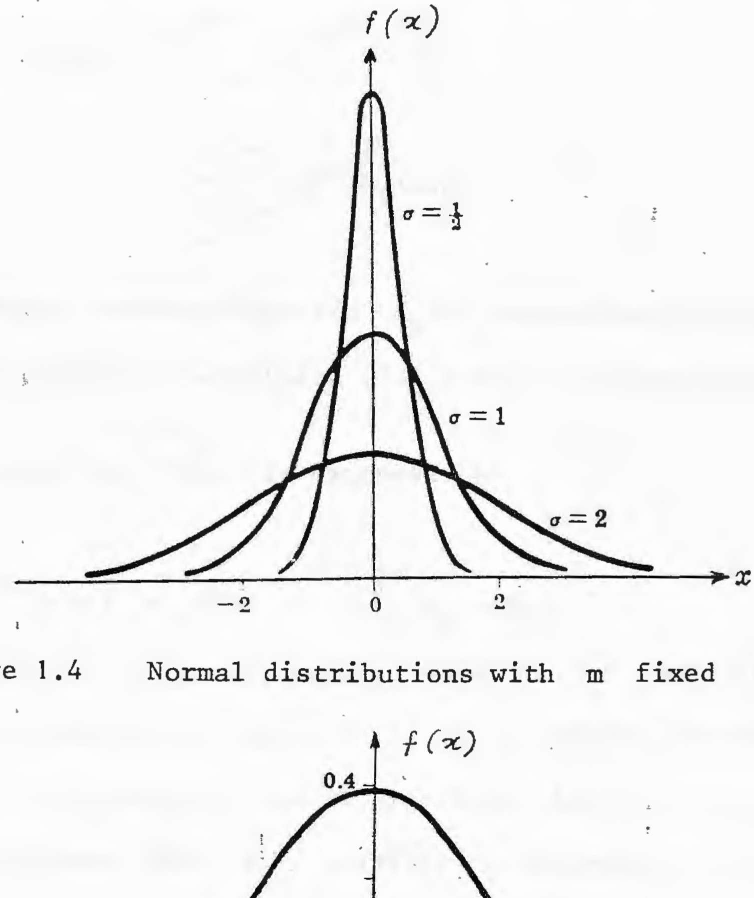


Figure 1.4 Normal distributions with m fixed (m = 0)

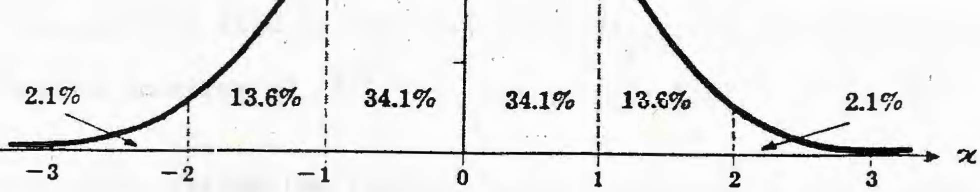


Figure 1.5 Normal distribution N (0,1) (m = 0, σ = 1)

For $-3 \leq x \leq 3$ we obtain 99.74% of the area under the curve.

A Gaussian process is completely characterized by its mean value and covariance. For a stationary Gaussian process, the mean value is constant and the covariance depends on $t_1 - t_2$ only.

As mentioned above, in a weakly stationary process $\{X(t), t \in T\}$ the mean value $m(t)$ and covariance $r(t)$ are constant. The spectral density function $S_X(\omega)$ of such a process can be defined from its covariance as:

$$S_X(\omega) = \int_{-\infty}^{\infty} e^{-j\omega t} r(t)dt \tag{1.53}$$

and

$$r(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega t} S_X(\omega)d\omega \tag{1.54}$$

The spectral density function $S_X(\omega)$ shows how the energy of the stochastic process is distributed with respect to frequency ω.

The variance of $X(t)$ is expressed as

$$\sigma^2(t) = r(0) = \frac{1}{\pi}\int_{0}^{\infty} S_X(\omega)d\omega \tag{1.55}$$

Now, if a stochastic input $f(t)$ to the control system is assumed to be stationary and Gaussian, and an expression for the power-density spectrum $S_f(\omega)$ is available, then it is possible to determine a suitable value for D defined from (1.3).

First, we will find an expression for $S_{f^{(1)}}(\omega)$ because we need the power-density spectrum of $f^{(1)}(t)$ and not of $f(t)$.

From (1.4), taking the Laplace transform (assuming zero initial conditions), we have:

$$F^{(1)}(s) = sF(s) \tag{1.56}$$

or $$|F^{(1)}(j\omega)|^2 = \omega^2|F(j\omega)|^2 \qquad (1.57)$$

Thus $$S_{f^{(1)}}(\omega) = \omega^2 S_f(\omega) \qquad (1.58)$$

and the variance σ^2 of $f^{(1)}$ can be obtained from the power-density spectrum of $f(t)$ using (1.55) and (1.48), i.e.

$$\sigma^2_{f^{(1)}}(t) = \frac{1}{\pi}\int_0^\infty \omega^2 S_f(\omega)d\omega \leq \infty \qquad (1.59)$$

and the standard deviation of $f^{(1)}$ is σ.

It has been found that the probability of $f^{(1)}$ being within $\pm 3\sigma$ is 99.74% or 0.9974 [8]. After this result, without a great mistake, the maximum magnitude of $f^{(1)}$ is 3σ, i.e.

$$D = 3\sigma \qquad (1.60)$$

This assumption is widely used and is often valid.

Chapter 2

HIGHWAY AUTOMATIC CAR-FOLLOWING CONTROL

2.1 Ground Transportation Systems

During the last 15 years, there has been a great interest in new classes of ground transportation systems. Most of this interest has dealt with system concepts and various related theoretical developments, and very little has been concerned with system development. A detailed analysis of such systems, especially of those dealing with vehicle control, is based upon idealised models, in which perfect information and control of the proposed vehicle types are assumed.

Several types of ground transportation systems have been suggested for future implementations. All these systems attempt to combine some of the desirable features of the private automobile (e.g. convenience, flexibility, comfort, speed) with the traditional advantages of public transportation systems, such as economy of travel, good safety record, relief from driving strain, etc.

Although proposed automated ground transportation concepts involve different engineering problems and solutions, their final feasibility often reduces to a common class of control technology problems. Difficult control and communication problems arise from the use of short spacings, the large number of vehicles and vehicle interaction in merging flows.

The main control functions - common in varying degrees to all automated ground transport systems - are the following, in brief:

1. Automatic Longitudinal Control.

It consists of either the control of a car with respect to the nearest lead car, or the control of an n-car traffic stream, for obtaining a maximum capacity of the system in a single-lane straight roadway.

2. Automatic Lateral (or Steering) Control.

It consists of the design of a suitable roadway reference for guidance, and the design of appropriate sensors, so that the position of the vehicle relative to the reference can be determined.

3. Automatic Vehicle-Spacing Detection

This detection contains a spacing policy that can be undertaken in the longitudinal control, to avoid collisions and reach great capacities of the system.

4. System Decision-Making Capability

This capability can be handled for all control decisions regarding the vehicles on a given section of highway, which would be made either by a centrally located computer, or by a small computer, or by both individual vehicles and centrally located computers.

5. Vehicle Communication Systems.

These must be such that the following vital functions are performed:

(a) communication with the driver,

(b) communication between the information collecting devices such as spacing detectors, lateral deviation measuring units, etc., and the decision-making systems or control systems,

(c) communication between the decision-making capability of the system and the vehicle control devices.

6. Automatic Merging Control

A vehicle merges when an acceptable gap appears in the traffic stream. An acceptable gap is defined as the time interval between the arrivals of two consecutive vehicles at a specified point. The mechanics of the merging manoeuver is dependent on the choices made for the functions of all the above mentioned systems 1 to 5. This dependence is a complex one because of the interrelationships existing among these functions.

7. Vehicle Propulsion Control

There are two general types of vehicle propulsion, one in which a vehicle is self-powered and another wherein the vehicle is self-powered on non-automated roads and externally powered on automated ones. The problems involved in making this choice are mainly nontechnical, such as air pollution, noise pollution, etc., and they have a considerable effect on the general community.

8. Automatic Vehicle Checkout

In order to merge a vehicle on to an automatic highway, it could be required to make automatically some manoeuvers before being permitted to join the traffic stream. (The checkout would require a minimum of time so as not to cause unnecessary delays in waiting and overall travel time.)

9. Automatic Routing and Dispatching Control

The usefulness of this function is to move the vehicles from the points of origin through the network to the preselected destinations along the shortest available paths.

10. Control in Stations

This control includes the stopping, unloading, loading and starting operations, as well as the in-station movements of vehicle queues.

11. Compatible Manual Mode Systems

It is necessary for a system to have manual control mode so that a driver can control the vehicle on non-automotive roads. This mode can be either conventional control, such as steering wheel, brake pedal and accelerator pedal, or some other type of manual control. The use of conventional controls would make the overall automated system more acceptable to the public. However, it is difficult to achieve compatibility between this control and an automated system.

12. Emergency Control

This function is used in order to override all other control functions and perform a prescribed operation in the event of a failure.

As mentioned above, the criteria for control are to achieve the desired quality of service, which can be expressed in prescribed travel time, safety level, ride quality, etc.

2.2 Introduction to Car-Following Control

The rapid growth of the world's population, the increased dependence on the automobile and the change from public to private transportation have contributed to increasing traffic congestion and a steadily rising highway accident rate.

If one would look further ahead to the end of the century, he would see vast sprawling supercities with populations characterised by adequate incomes, longer life spans, and increased amount of leisure time. One predictable result is greatly increased travel. The resulting travel situation will be chaotic, unless some radical changes are undertaken.

One partial solution to such problems involves the construction of more highways; however, the resulting enormous cost is certain to limit such constructions. Consequently, it is necessary to examine other methods of safely and efficiently handling the expected traffic problem.

Any approach taken should increase both highway capacity and highway safety. The increase in capacity under conditions of high-density high-speed traffic flow can be obtained by reducing the distance between the vehicles in a moving traffic stream. However, the driver can do relatively little to help attain this goal because of both his slow reaction time and limited visual acuity.

One attractive partial solution is highway automation. It is useful first to distinguish between single-mode systems with automatically controlled captive-vehicles on a guideway network (special roadway), and dual mode systems, wherein the individual vehicles are automatically operated on the guideway, but the driver resumes manual control after leaving the guideway.

The advantages of the captive-vehicle systems include transportation, to all citizens in a restricted geographical area, like a downtown business district, and the partial or complete elimination of private vehicles from that area and besides a relief from typical urban transportation problems such as air pollution, noise, congestion, etc.

A dual-mode system offers the prospects of an increased traffic flow and improved safety - at least on the automated part of the system - by replacement of the human driver with a faster reacting and more reliable automatic control system.

For achieving a highway automation, the most frequently suggested system involves a roadway complex, consisting of both automated and non-automated roads. The main highways would be equipped for automation, and various rural roads and urban streets would not be equipped. This condition would not come into existence overnight.

In broad outline, an individual vehicle would enter the system at a special entrance point, where it would first undergo a rapid automatic checkout, and the driver would indicate his destination. If the vehicle passed the checkout, it would be automatically merged into the traffic stream.

The vehicle would remain under automatic control until the driver's preselected exit was reached; then it would be guided off the highway onto the exit ramp and the control would be returned to the driver.

One can expect two primary benefits from this system: (i) greatly increased lane capacity, which would depend, of course on the chosen system design, and (ii) a reduction in the number of highway accidents. The expectation of fewer accidents arises from both the rapid reaction time and consistency of an electronic system in comparison with a driver.

The present work is based on the fact that an automatic system is gradually introduced which is compatible with existing traffic at all stages.

The resulting design (based on Zakian's theory) will, however, be only for longitudinal control and specially for the case of a steady-state car-following mode.

2.3 System Requirements

In order to obtain any satisfactory controller in a steady-state car-following mode, there are some assumptions and requirements which are associated with safety considerations, passenger comfort, traffic density and fuel economy. These restrictions that must be met are the following:

1. All vehicles in the string have identical dynamics.
2. All vehicles are operating on a single lane of straight and level roadway.
3. The average separation between adjacent vehicles must not be excessive; the smaller the spacing at a fixed velocity, the greater the capacity of the system.
4. Disturbances propagated to the rear must be attenuated or at least limited in amplitude (asymptotic stability).
5. These disturbances affecting the motion of the vehicles (winds, gusts, etc.) are random (zero-mean Gaussian white noise) in nature.

6. To provide a smooth economical ride and ensure the comfort of the occupants, the normal acceleration and deceleration of a vehicle would normally not exceed 0.1 g.

7. The control system must not be required to respond so as to exceed the response capabilities of the vehicle.

The headway h is defined as the distance between the rear end of the lead vehicle and the front end of the following vehicle (to be distinguished from vehicle spacing, which also includes the car length ℓ).

If the minimum headway is denoted by h_o, then for a given average stream velocity V_s the line capacity C is:

$$C = \frac{V_s}{h_o + \ell} \quad \text{(cars/lane)/sec} \tag{2.1}$$

It is clear that if speed and headway were not related, the line capacity could be increased, either by increasing the main line velocity, or by reducing the minimum headway requirement. The main line velocity, however, is related to the minimum headway through its effect upon the safe stopping distance. There is a relation between the relative velocity v and headway h,

$$v = \frac{dh}{dt} \tag{2.2}$$

It follows from these requirements that necessary inputs to the longitudinal control system are the instantaneous headway and relative velocity existing between the controlled and lead vehicle. These quantities are easily obtained in an experimental situation by using a mechanical take-up reel in conjunction with a wire strung between two cars; however, this is impractical for general use. A number of researchers have examined this problem and suggested various techniques for obtaining these quantities, such as the use of radar, infra red communication, or electronic "block" system mounted in the pavement.

2.4 System Description

It has been shown [13] that the above requirements can be met by using a multimode control system, which can be conveniently described with a two-dimensional phase plane. Here the relative velocity v† between two vehicles is plotted versus the headway h between those vehicles. Time is not shown explicitly on this plane, but as time progresses, the point representing the instantaneous relative velocity and headway traces a trajectory on the plane.

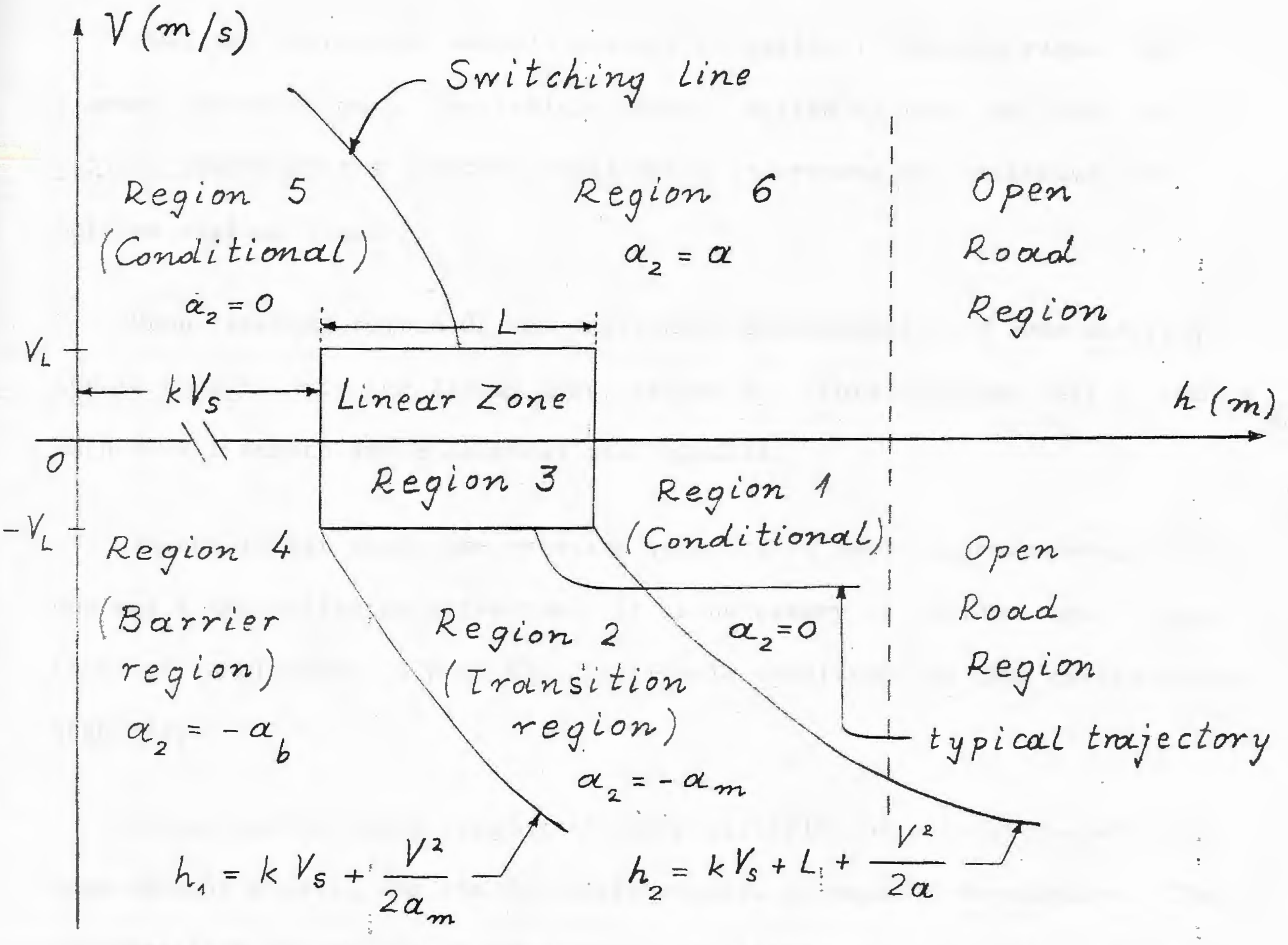


Figure 2.1 Regions of phase plane

Consider the phase plane shown in Figure 2.1 . It is divided into a number of regions, a certain mode of control being associated with each region. These are separated by switching boundaries, which are defined in the figure.

† $v = v_1 - v_2$ (v_1 is the velocity of the lead vehicle and v_2 is the velocity of the following vehicle).

Thus, as a trajectory moves and crosses a boundary, the control mode changes. This change is made by a simple logic system, which associates each point in the phase plane with a particular mode of control.

When the controlled vehicle is in the right-most region of the phase plane (open-road region), the control system behaves as a velocity regulator with the command velocity set either by the driver or a traffic controller. Here the headway is sufficiently large that the two vehicles are "uncoupled".

When the controlled vehicle moves into Region 1 from the right, the command acceleration to the vehicle control system is zero, and thus the vehicle proceeds at a constant speed until it crosses the switching line between regions 1 and 2.

Upon entering region 2, the vehicle is decelerated at a constant rate and is brought into the linear zone (region 3). This constant rate is chosen such that a smooth and economical ride results.

In the linear zone, the relative velocity is approximately zero, and one has a car-following situation. It is necessary to include such a region in order to eliminate a possible limit-cycle condition and thus obtain system stability.

In the barrier zone (region 4) the possibility of a collision with the lead vehicle exists, and the following vehicle is rapidly decelerated. The distance from the origin to the edge of the linear zone is equal to the time headway k times the average stream speed V_s. The parameter k is determined by safety considerations and is specified in terms of lead-car average speed and the deceleration capabilities of the lead and controlled vehicles.

This factor k, together with the stream speed V_s, determines the vehicle flow rate. The basic equation of traffic flow is

$$\rho = \lambda V_s \qquad (2.3)$$

where ρ is the capacity (cars/lane)/hr, λ is the concentration in (cars/lane)/km, and V_s is expressed in km/hr. It has been found for the case under consideration, that for the maximum flow λ must be:

$$\lambda = \frac{3600}{L + k V_s} \tag{2.4}$$

and equation (2.3) becomes:

$$\rho = \frac{3600}{(L/V_s) + k} \tag{2.5}$$

where L is the length of the linear zone.

Thus the maximum flow capacity ρ of the system is fixed by k and the length L of the linear zone, so k must be kept small if one is to achieve high flow capacities.

Now the control region 5 is dependent upon how this region is entered from the linear zone, the controlled vehicle is accelerated at a rate of +0.1g, so as to close the "gap" before it becomes too large. However, if it is entered from the barrier zone (where there is the possibility of a collision), the controlled vehicle is allowed to coast, which will ensure that it will soon be in region 6. This is an acceleration region which can be entered from either region 5 or the linear zone.

For example, if the two vehicles are in a steady state, and the lead vehicle is suddenly accelerated, the resultant state will be "in" region 6. The magnitude and duration of the lead vehicle acceleration will determine whether the vehicles are decoupled or returned to a steady-state car-following situation. A number of other choices are possible.

It is useful to consider a simple example, namely the case where a fast moving vehicle is overtaking a slower vehicle. It is assumed that the initial speeds of both vehicles are constant. The resulting trajectory (typical trajectory), shown in Figure 2.1, consists of a horizontal line up to the point where a threshold curve is "crossed". At this point the vehicle is decelerated at a constant rate and "directed" toward the linear zone. The vehicle, once "in" the linear zone, will tend to remain there unless a disturbance is introduced into the system.

2.5 Linear Velocity Control

Consider a line of identical vehicles that are attempting to follow one another in a stable manner. For simplicity, only the first two vehicles in the line are considered; however, the results obtained also apply to the ith and the (i+1)th vehicles. It is assumed that the vehicles are initially in a steady-state condition, i.e. they are in the linear zone.

The control law chosen for the linear mode of operation is

$$\frac{dv_2}{dt} = \frac{1}{\tau} v = \frac{1}{\tau}(v_1 - v_2) \tag{2.6}$$

where dv_2/dt is the command acceleration of controlled car

v_1 is the variational component of lead-car speed

v_2 is the variational component of controlled-car speed

τ is small-signal time constant of controlled car.

It must be noted that this law was first studied by Pipes [28] as a model for single-lane traffic flow. Subsequently, researchers introduced driver response time into this equation and showed that it gave a good approximation to the car-following performance of the driver vehicle unit [9].

Now we consider the transient behaviour of the controlled vehicle after it enters the linear-control zone at a point A, as shown in Figure 2.2 .

If the speed $v_1(t)$ of the lead car is assumed constant at V_1 then, solving the differential equation (2.6) for v_2, we get the speed of the controlled vehicle:

$$v_2(t) = V_1 + V_L \exp(-\frac{t}{\tau}) \tag{2.7}$$

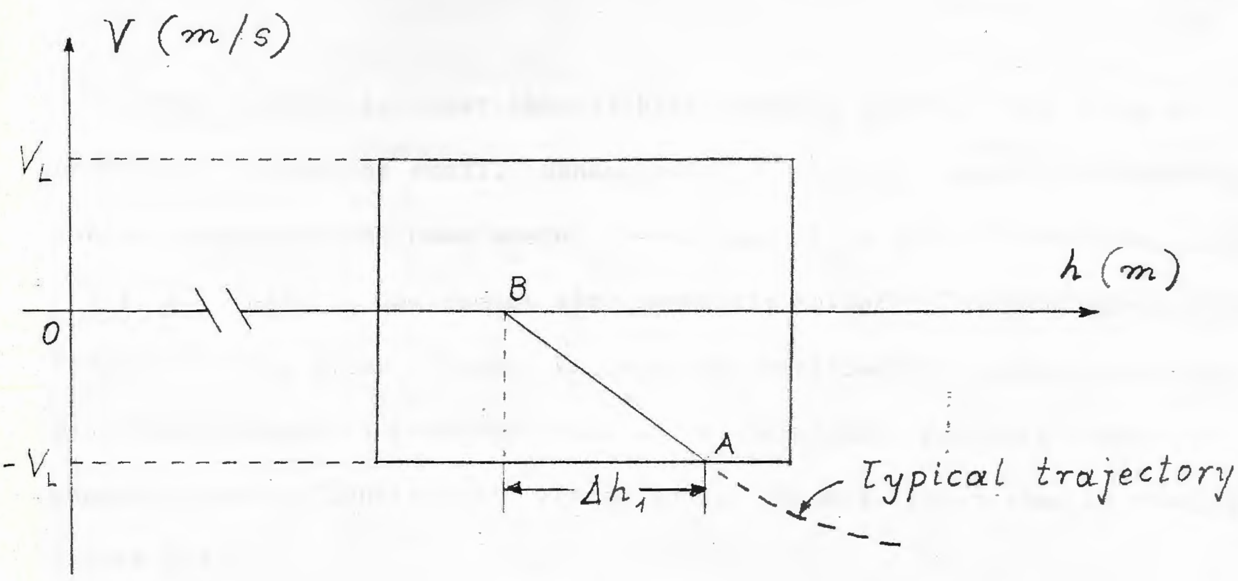


Figure 2.2 Transient behaviour of controlled car on the v-h plane

Here V_L is the relative velocity at the linear-zone boundary. The total change Δh_1 in headway before the controlled vehicle is in equilibrium at point B with respect to the lead car, is:

$$\Delta h_1 = \int_0^\infty v(t)dt = -\tau V_L \quad (2.8)$$

as can be determined from (2.2) . Thus, the change in headway is determined by the product of the system time constant and the relative velocity on the lower boundary of the linear zone.

As is clear from (2.8), a constraint is posed on the length L of the linear zone

$$L \geq |\Delta h_1| = \tau |V_L| \quad (2.9)$$

A stronger condition which one would try to meet in practice in order to prevent the overtaking vehicle from being frequently taken into the barrier region is:

$$L \geq 2\tau |V_L| \quad (2.10)$$

From (2.5) it is clear that if high capacity traffic flow is to be achieved, L must be small. Hence, both τ and V_L should be minimised. Taking as example the case where $\tau = 4s$ and $V_L = 2m/s$, then from (2.10) $L \geq 16$ m. Clearly one cannot simultaneously satisfy (2.9) and achieve high-density traffic flow. Thus, V_L must be specified in a manner such that passenger comfort is preserved and L is minimised. For this reason, we consider next a "conditional" linear zone. Such a linear zone is shown in Figure 2.3 .

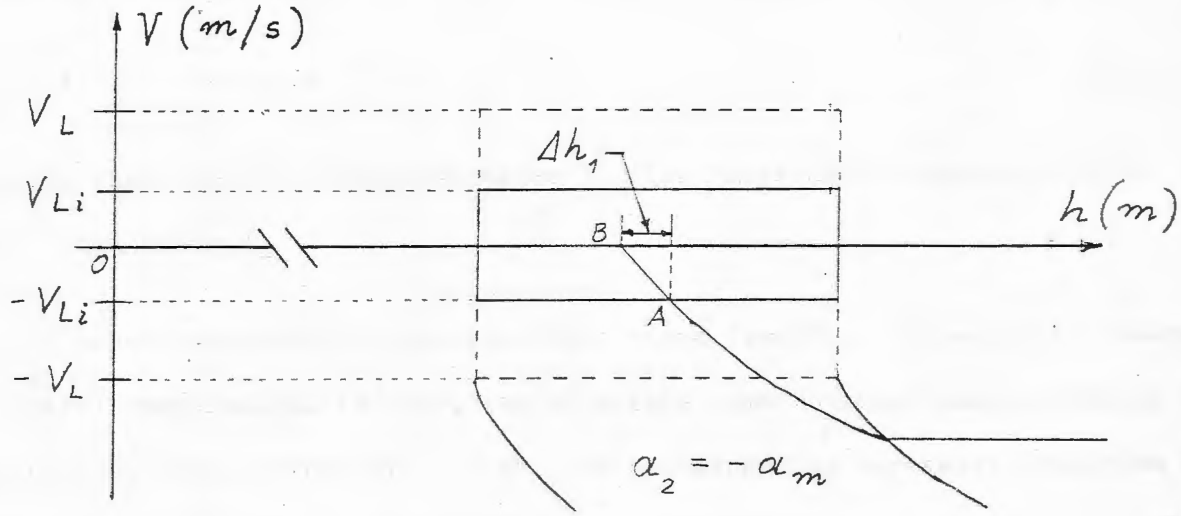


Figure 2.3 v–h plane with conditional zone.

In this configuration, the upper and lower velocity boundaries of the linear zone are positioned at a low value of relative velocity V_{L_i} when the controlled vehicle is not in this zone. However, after the vehicle enters the linear zone, either from above or below, the velocity boundaries are moved to a larger preselected value V_L. This preselected value can be specified by (2.9) or other conditions which are associated with safety and passenger comfort. The change in headway after one enters the linear zone is now given by:

$$\Delta h_1 = \tau |V_{L_i}| \tag{2.11}$$

Thus, for $\tau = 4s$ and $V_{L_i} = 0.305$ m/s, then $L \geq 2.44$ m.

In fact, this effective time constant τ is the controlled vehicle response to small changes in command velocity and is defined as the time required for the vehicle to reach 0.632 of its final velocity value for a command step change in velocity.

It has been found [5] that a necessary condition for asymptotic stability is:

$$k \geq 0.787\tau \qquad (2.12)$$

A stronger condition is

$$k \geq \tau \qquad (2.13)$$

However, there are circumstances where a value intermediate between 0.787τ and τ can be used.

In practice the usual time constants range from 8 to 30 seconds. However, if control compensation is used, the effective time constant can be reduced to 4 seconds or less. Even for $r = 4s$, the corresponding necessary condition is $k \geq 3.148$ and one has a relatively large value of the time headway k. If one is to have both asymptotic stability and a small headway $(k < 1)$, it would be necessary to reduce τ to 1 second or less.

2.6 Model Analysis

The longitudinal dynamics of the following vehicle for small signal case [5] is given by a gross approximated transfer function:

$$G(s) = \frac{V_2(s)}{T_v(s)} = \frac{K}{Ts + 1} \qquad (2.14)$$

where T_v represents the incremental change in throttle valve position from a steady-state position T_{vo}, T equals the uncompensated vehicle time constant, and K equals the steady-state gain from throttle valve to output speed. It should be noted that the dynamics of the electrohydraulic throttle valve control system were negligible in comparison with vehicle dynamics and hence were not considered.

Todoslev [36] and Blackwell [7] have shown that this representation is valid under steady-state small signal conditions.

K and T are, in practice, functions of speed and a variety of environmental effects. Bender [4] reported that T varied from 5 to 40 sec for small variations ± 1.5 m/sec about an average speed of 22.3 m/sec (80.3 km/h) on a nearly straight and level highway, and K varied from 1 to 6 .

Such variations can be overcome by using internal velocity feedback through a gain of $f = 8$ as shown in Figure 2.4 .

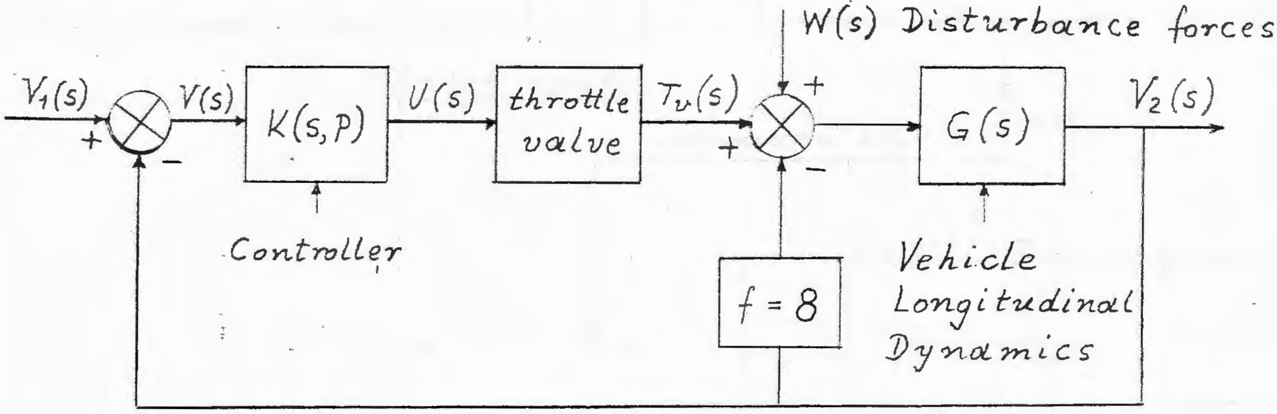


Figure 2.4 Vehicle velocity controller with internal velocity feedback (linear mode).

Now, with this parameter f introduced the effective vehicle dynamics become:

$$G_1(s) = \frac{G(s)}{1 + f\,G(s)} = \frac{K_1}{T_1 s + 1} \qquad (2.15)$$

where

$$K_1 = \frac{K}{1+fK} \quad \text{and} \quad T_1 = \frac{T}{1+fK} \qquad (2.16)$$

and the reduced diagram is shown in Figure 2.5 .

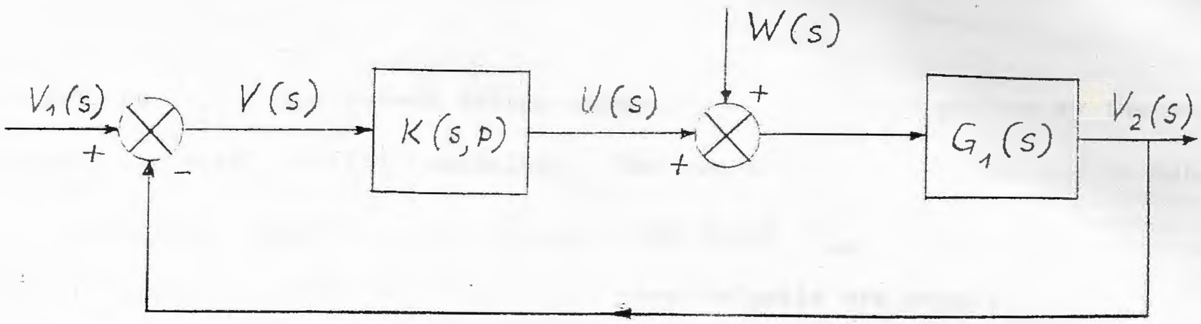


Figure 2.5 Reduced diagram (linear model)

A block diagram of the overall system configuration is shown in Figure 2.6, where $G'(s)$ represents the effective compensative dynamics.

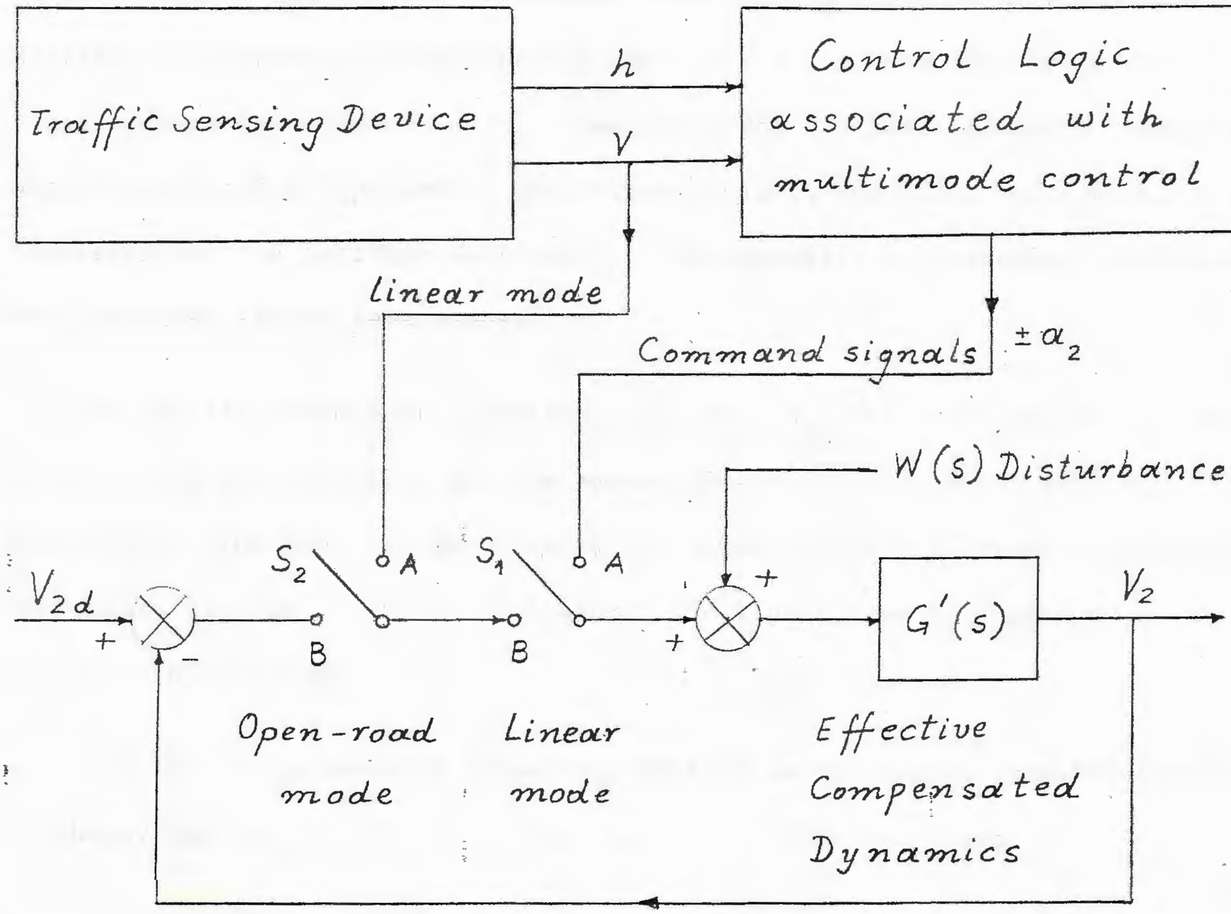


Figure 2.6 Overall system configuration.

The control logic responds to the sensed traffic variables by properly positioning switches S_1 and S_2. When the controlled vehicle is in the right-most region of the phase plane shown in Figure 2.1, i.e. in the open road region, both switches S_1 and S_2 are in the B position and the system input is V_{2d} - a constant driver command speed selected either by the driver or by a central traffic controller. The resulting closed-loop system behaves as a velocity regulator and maintains the speed V_{2d} . Note that under all other traffic conditions, the system input signals are supplied by either a traffic sensing device or from circuitry associated with the control logic.

When the controlled vehicle moves into region 1 from the right, S_1 is moved to the A position and the input to the effective vehicle plant is zero. Thus, the vehicle continues to maintain the preselected speed V_{2d} . However, if this region were entered from above, the vehicle would be accelerated to decrease the headway as described in Section 2.4 . When the controlled vehicle moves into region 2, S_1 remains in the A position and a constant negative voltage is applied to the integral plant, resulting in a constant deceleration. (A positive acceleration corresponding to operation in region 6 is obtained in the same manner.)

In the linear mode of operation, S_1 and S_2 are moved to the B and A positions respectively, and the system functions as a linear velocity controller. The logic is constructed such that conditional relative velocity boundaries are used, and the position of the linear zone is dependent on the control vehicle speed.

For the linear mode of operation, we take as an overall transfer function of the system (as considered in the case study [32]) the equation:

$$G_1(s) = \frac{V_1(s)}{V_2(s)} = \frac{1}{4s+1} \tag{2.17}$$

where $V_1(s)$ and $V_2(s)$ are the Laplace transforms of $v_1(t)$ and $v_2(t)$ respectively. This transfer function comes from the following process.

The small signal car-following model is such that the incremental acceleration dv_2/dt of the controlled vehicle is equal to a linear combination of relative velocity v and a headway error term. Thus:

$$\frac{dv_2}{dt} = k_1 v + k_2(h-k_3 v_1-k_4 v_2) \tag{2.18}$$

where v_1 is the incremental speed change of the lead car, h is the incremental change in headway between the lead and controlled vehicles, and k_i, $i = 1,2,3,4$, are constants with $k = k_3+k_4$. Generally, either k_3 or k_4 is zero.

Having Laplace transforms of (2.18) with zero initial conditions and taking into account that

$$v = v_1 - v_2 = \frac{dh}{dt} \tag{2.19}$$

we find

$$sV_2(s) = k_1[V_1(s)-V_2(s)] + k_2\left[\frac{V_1(s)-V_2(s)}{s} - k_3V_1(s) - k_4V_2(s)\right]$$

or, after some rearranging,

$$\frac{V_2(s)}{V_1(s)} = \frac{(k_1-k_2k_3)s + k_2}{s^2+(k_1+k_2k_4)s+k_2} \tag{2.20}$$

For a locally stable system it is necessary that the coefficients of the denominator polynomial be positive, i.e.

$$k_2 > 0, \quad k_1+k_2k_4 > 0 \tag{2.21}$$

Several different car-following situations were examined. One of them was that where

$$k_1 = 0.25, \quad k_2 = 0.5, \quad k_3 = 0 \text{ and } k_4 = 4 \tag{2.22}$$

so that equation (2.18) becomes:

$$\frac{dv_2}{dt} = 0.25v + 0.5(h-4v_2) \tag{2.23}$$

Thus, the system transfer function (2.20), after simplifications, takes on the form of equation (2.17).

2.7 Determination of the Input Space

The main disturbances W(s) for a car-following system are aerodynamic drag forces caused principally by winds and gusts, as well as by the vehicle's nominant speed. An approximate expression of an aerodynamic drag force [31] or [24] is given by:

$$F_D = \frac{1}{2}\rho V_o A V_w C_D + R \tag{2.24}$$

where $\frac{1}{2} V_o A V_w C_D$ is the actual aerodynamic drag and

R = rolling resistance

ρ = mass density of air

V_o = nominal speed of vehicle

A = projected frontal area of the test vehicle

V_w = wind speed with respect to the ground

C_D = drag coefficient.

Variations in the wind speed cause variations in the car speed, as can be seen from (2.24), which can be positive or negative according to the direction of the wind speed. (For simplicity, only the longitudinal wind speeds are considered.)

As discussed previously (Section 1.7B), it is possible to estimate the number D, defined from (1.3), if the input to the system is a stationary and Gaussian, and an expression for the power density spectrum is available.

It has been suggested [26] a power density spectrum $S_w(\omega)$ of the wind forces of the form:

$$S_w(\omega) = \frac{c^2}{1 + \omega^2/d^2} \tag{2.25}$$

where c is a parameter dependent on the form of the front of the vehicle, and the aerodynamical drag forces, and d is a constant dependent on the geographical area of the highway. The disturbance is assumed to be zero mean Gaussian white noise, as emphasized in Section 2.3, but wind as a stochastic process is stationary only when considered over short periods of time. Thus, the properties of the disturbances will be constantly changing and the power density spectrum will not necessarily be valid for a long time. So the values for c and d obtained are not always valid. It has been found for wind speeds up to 80 km/h, for nominal vehicle speeds up to 90km/h and a frontal area of the vehicle 0.7 m^2 in open highways without any mountains, hills, etc., that c and d take on the worst values:

$$c = 0.326 \quad \text{and} \quad d = 0.5 \text{ rad/sec} \tag{2.26}$$

Now from Figure 2.5 we have:

$$V(s) = V_1(s) - V_2(s) \tag{2.27}$$

and

$$V_2(s) = G_1(s)\{W(s) + K(s,p)V(s)\} \tag{2.28}$$

These last equations give:

$$V(s) = \frac{V_1(s) - G_1(s)W(s)}{1+K(s,p)G_1(s)} \tag{2.29}$$

The latter equation compared with equation (1.8) gives as a total input to the system,

$$F(s) = V_1(s) - F_d(s) \tag{2.30}$$

where $$F_d(s) = G_1(s)W(s) \tag{2.31}$$

and the block diagram of Figure 2.5 takes on the form of Figure 2.7 .

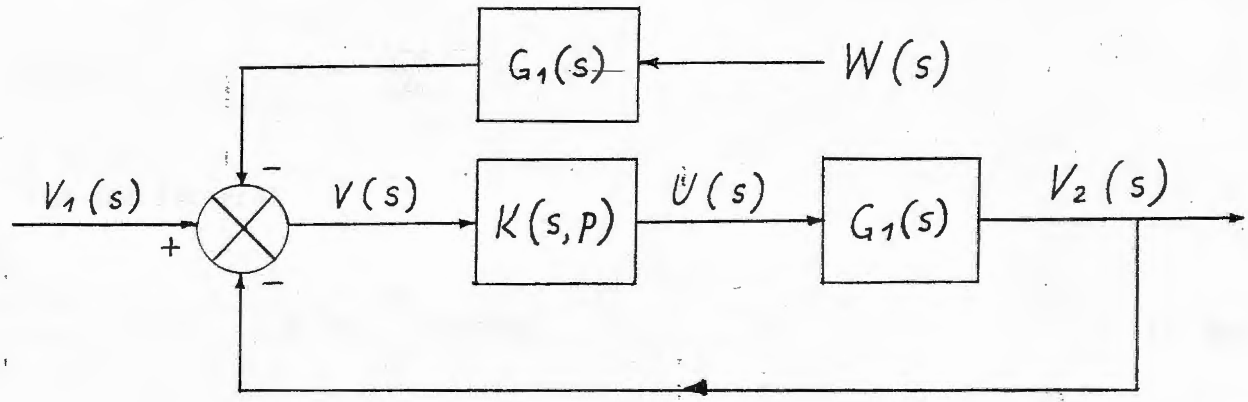


Figure 2.7 Block diagram of the car-following system with inputs (linear mode)

From equation (2.31) taking frequency responses

$$|F_d(j\omega)|^2 = |G_1(j\omega)|^2 |W(j\omega)|^2 \tag{2.32}$$

But $|W(j\omega)|^2$ is the power density spectrum $S_W(\omega)$. Hence,

$$|F_d(j\omega)|^2 = \frac{1}{1+16\omega^2} \cdot \frac{0.1063}{1+4\omega^2} \tag{2.33}$$

Now, from the power density spectrum of $F_d(s)$ using equation (1.56), we have

$$|F_d^{(1)}(j\omega)|^2 = \omega^2 |F_d(j\omega)|^2$$

$$= \frac{0.1063\omega^2}{(1+16\omega^2)(1+4\omega^2)} \tag{2.34}$$

and the variance σ^2 of $f^{(1)}(t)$ can be obtained using equation (1.55),

$$\sigma^2 = \frac{1}{\pi}\int_0^\infty \frac{0.1063\omega^2 d\omega}{(1+16\omega^2)(1+4\omega^2)} \tag{2.35}$$

Integrating numerically using Simpson's rule:

$$I = \frac{h}{3}\{y_o + \sum_{i=1}^{m} (4y_{2i-1}+2y_{2i}) + y_{2m}\} \tag{2.36}$$

where $h = \frac{b-a}{2m}$ (2.37)

for the integral

$$I = \int_a^b f(x)dx \tag{2.38}$$

we find $\sigma^2 = 8.036266\cdot 10^{-4}$ (2.39)

so that $\sigma = 0.028348$ (2.40)

and $D = 3\sigma = 0.085045$ (2.41)

Using a computer we find the plot of the power density spectrum of $F_d^{(1)}$, which is plotted in Figure 2.8 .

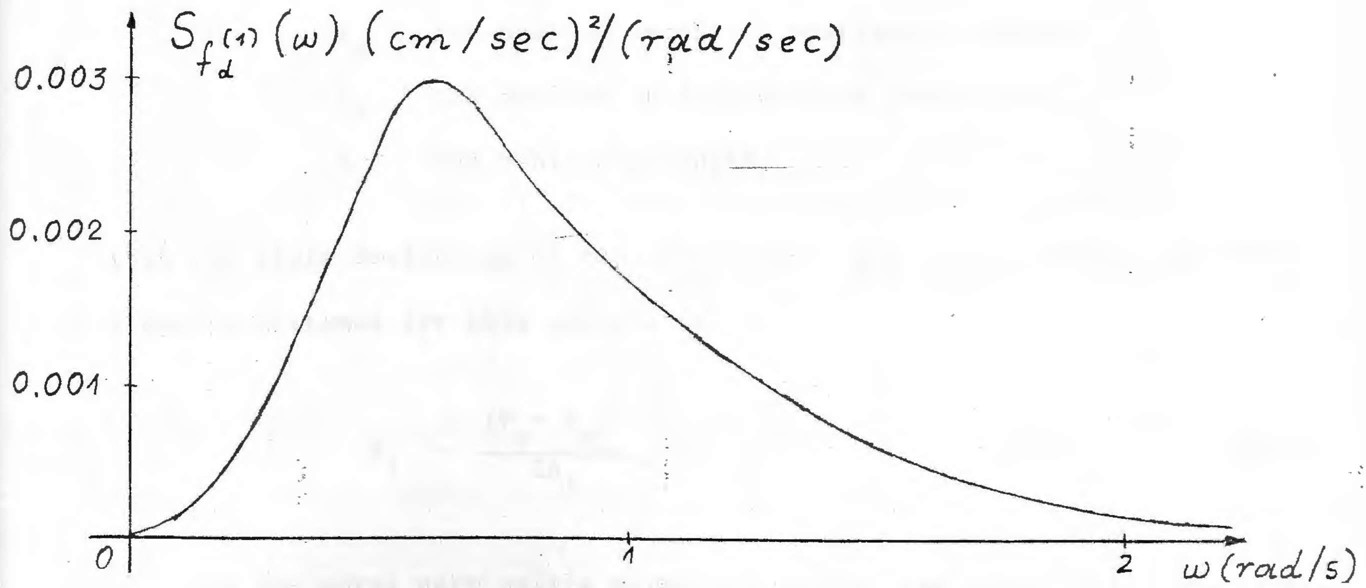


Figure 2.8 Power density spectrum of $F_d^{(1)}$

2.8 Criteria Specifications

The design problem focuses on finding a useful measure of performance, while the control problem is concerned with achieving a desired level of performance.

In our case, the problem is to specify an upper bound ε, (the greatest tolerated relative velocity error), which satisfies $\phi(p) \leq \varepsilon$, where $\phi(p)$ is defined from (1.9), achieving the greatest possible capacity of the system and avoiding a possible collision. Thus, we must consider the "worst" permissible case in order to get a "good" value for ε .

A worst—case situation is defined as one wherein the lead vehicle when it is in a worst permissible normal state suddenly decelerates at a rate of A_1 m/sec^2. After a communication delay of no more than t_c sec, a computer at the wayside is notified of the change in this state of the vehicle. Subsequently, after a second delay t_p, due to computer processing and retransmission time, the following vehicle is commanded to brake.

If we define with:

V_s	the desired stream speed
ΔV_m	the maximum permitted speed deviation
ΔX_m	the maximum permitted position deviation
k_d	the desired minimum headway time in sec
ℓ	the vehicle's length,

then the state deviations of the lead vehicle are $(-\Delta X_m, -\Delta V_m)$ and the stopping distance for this vehicle is:

$$S_1 = \frac{(V_s - V_m)^2}{2A_1} \tag{2.42}$$

For the worst permissible normal situation, the state variations of the following vehicle are specified as $(+\Delta X_m, +\Delta V_m)$. At time $t = 0$ the initial separation $d(0)$ between it and the lead vehicle is

$$d(0) = V_s k_d - 2\Delta X_m - \ell \qquad (2.43)$$

The following vehicle receives a deceleration command at time $t = t_c + t_p$ and subsequently decelerates at a rate A_2 m/sec^2 as it is shown in Figure 2.9 .

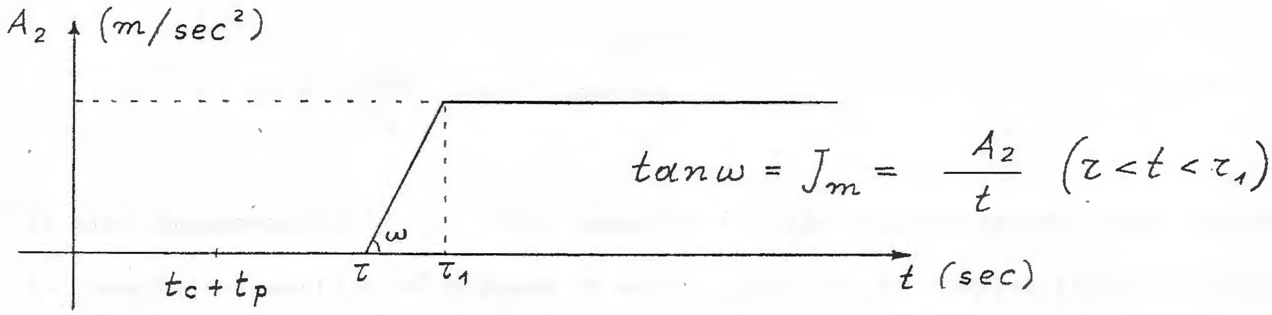


Figure 2.9 Deceleration trajectory of the following car

Note the presence of a braking reaction time τ, a finite jerk J_m and the selected deceleration A_2 . The stopping distance S_2 for the following car travelled from $t = 0$ is:

$$S_2 = (V_s + \Delta V_m)(T + \frac{A_2}{J_m}) - \frac{A_2^3}{6J_m^2} + \frac{(V_s + \Delta V_m - \frac{A_2^2}{2J_m})^2}{2A_2} \qquad (2.44)$$

where $T = t_c + t_p + \tau$ is the total reaction time delay.

To prevent a collision between the lead and the second vehicle it is necessary that

$$d(0) \geq S_2 - S_1 \qquad (2.45)$$

Substituting S_1 and S_2 from (2.42) and (2.44) respectively, into (2.45) and rearranging, there results:

$$V_s k_d \geq \ell + 2\Delta X_m + (V_s + \Delta V_m)(T + \frac{A_2}{J_m}) - \frac{A_2^3}{6J_m^2} + \frac{(V_s + \Delta V_m - \frac{A_2^2}{2J_m})^2}{2A_2} - \frac{(V_s - \Delta V_m)^2}{2A_1} \qquad (2.46)$$

Note from (2.46) that k_d is dependent on V_s. Thus, the maximum capacity C given by:

$$C = \frac{3600}{k_d} \text{ (cars/lane)/hr} \qquad (2.47)$$

is also dependent on V_s. This capacity is only achievable on a long linear (or nearly so) section of highway on which there are no complex traffic intersections.

There are seven parameters in (2.46) with two $(\Delta V_m, \Delta X_m)$ associated with the performance of the vehicle longitudinal controller, two (J_m and A_2) with the physical capabilities of the vehicle, two (T and k_d) which are largely associated with network operations, and one (A_1) which is a policy choice for the deceleration of the lead vehicle. All these parameters affect the minimum permissible headway.

The expected value or range of values for each of these can be specified considering such factors as the expected characteristics of future vehicles, the accuracy with which various state variables can be measured, and assumed achievable values of vehicle decelerations and deceleration rates (jerks).

So, [12], because of the trend towards smaller fuel-efficient vehicles, ℓ was selected as 3.67 m.

J_m was selected as 76.2 m/s^3 (250 ft/s^3) under the suppositions that this rate would be readily achievable in future vehicles and all vehicle occupants would be belt constrained in their seats.

Various industrial concerns have reported that a lower limit on τ is some 0.1 sec. Thus, this value was taken as a lower limit on the total delay time. The upper limit was selected as 0.6 sec because larger values would make the required capacity very difficult to obtain for any realistic choices of the other parameters.

The range of possible lead-vehicle and following-vehicle deceleration was selected to be 3.92 - 7.85 m/s^2 (0.4 g to 0.8 g) if the vehicles were not involved in a front-end collision. The highest value corresponds to braking at maximum abnormal situations.

The quantity ΔX_m is comprised of two components, one ΔX_a associated with the accuracy with which the vehicle's position could be measured, and a second ΔX_d with vehicle-position deviations under various manoeuvering and disturbance conditions. In practice, $\Delta X_m \leq 0.3048$ m (1 ft) appears to be achievable.

The quantity ΔV_m is the sum of ΔV_a - the accuracy to which instantaneous velocity can be measured, - and ΔV_d - the maximum deviation in instantaneous velocity expected under normal operating conditions. The most accurate reported technique for obtaining ΔV_a has an accuracy of ± 0.1524m/s (0.5 ft/s). The measurement involved includes a measurement time of some 0.05 - 0.1 sec and thus an effective delay is present. For a good vehicle controller

$$|\Delta V_d(t)| < 0.1524 \text{ m/sec} \qquad (2.48)$$

should be achievable. Thus, a limit on ΔV_m would be 0.3048 m/s (1 ft/s), i.e.

$$|\Delta V_m(t)| \leq 0.3048 \text{ m/sec} \qquad (2.49)$$

However, a practical choice would be:

$$0.3048 < \Delta V_m < 0.6096 \text{ m/s} . \tag{2.50}$$

This value of (2.49) can constitute the upper bound ε as a performance criterion for a velocity control in the linear mode of the car-following system and the value of (2.48) as a very tight bound ε .

The importance and effect of each parameter in (2.46) is different from the total system performance. For example C is relatively insensitive to changes in ℓ, J_m and ΔX_m, unless these changes are large, e.g. an increase of 1.8 m in vehicle's length ℓ. However, C is quite sensitive to even modest changes in T, ΔV_m, A_1 and A_2, which are characterised especially critical quantities.

The influence of T is shown in Figure 2.10, where V_s versus C is presented with T as a parameter and fixed $A_1 = 0.8g$, $A_2 = 0.6g$ and $\Delta V_m = 0.3048$ m/s. Note that a peak capacity of about 3,850 veh/lane/h was obtained for $T = 0.2$ at a stream velocity of about 50 km/h (13.9 m/s), whereas this capacity decreased to 2,675 veh/lane/h for $T = 0.6$ sec.

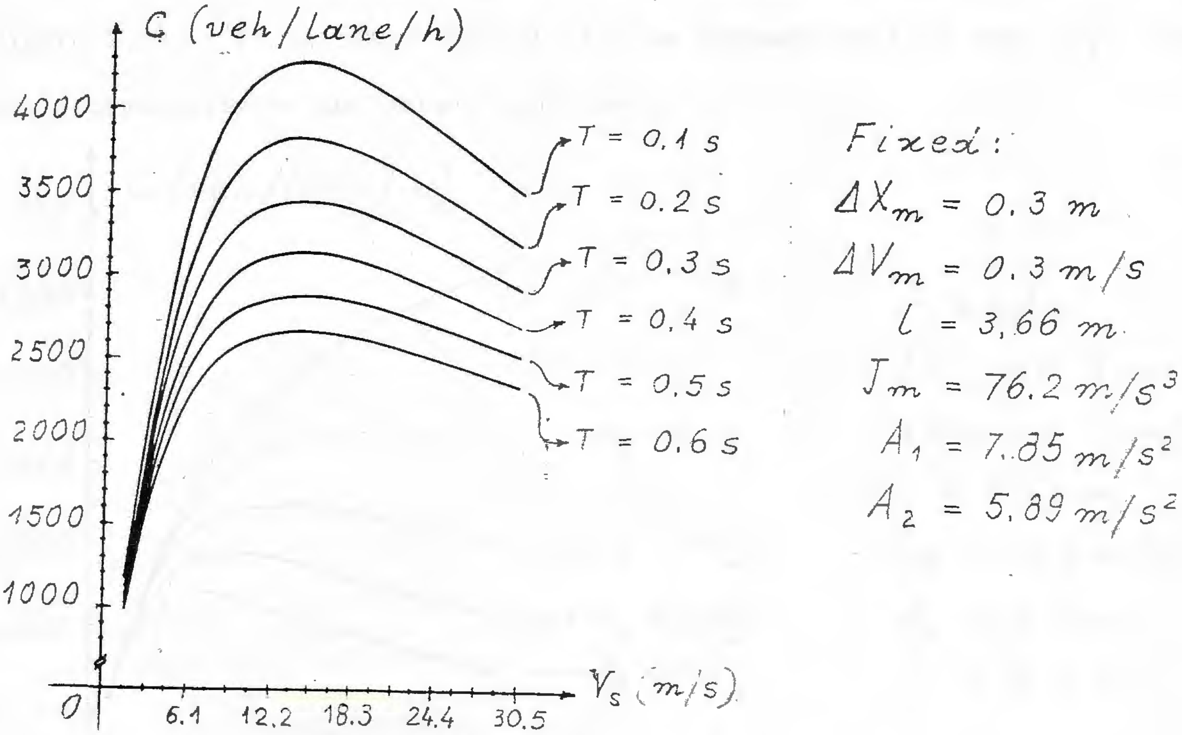


Figure 2.10 Maximum capacity C versus speed V_s with T as a parameter.

The effect of ΔV_m on capacity is shown in Figure 2.11, where V_s versus C is presented with ΔV_m as a parameter and fixed $A_1 = 0.8g$, $A_2 = 0.6g$ and $T = 0.3$ sec. Obviously ΔV_m must be small.

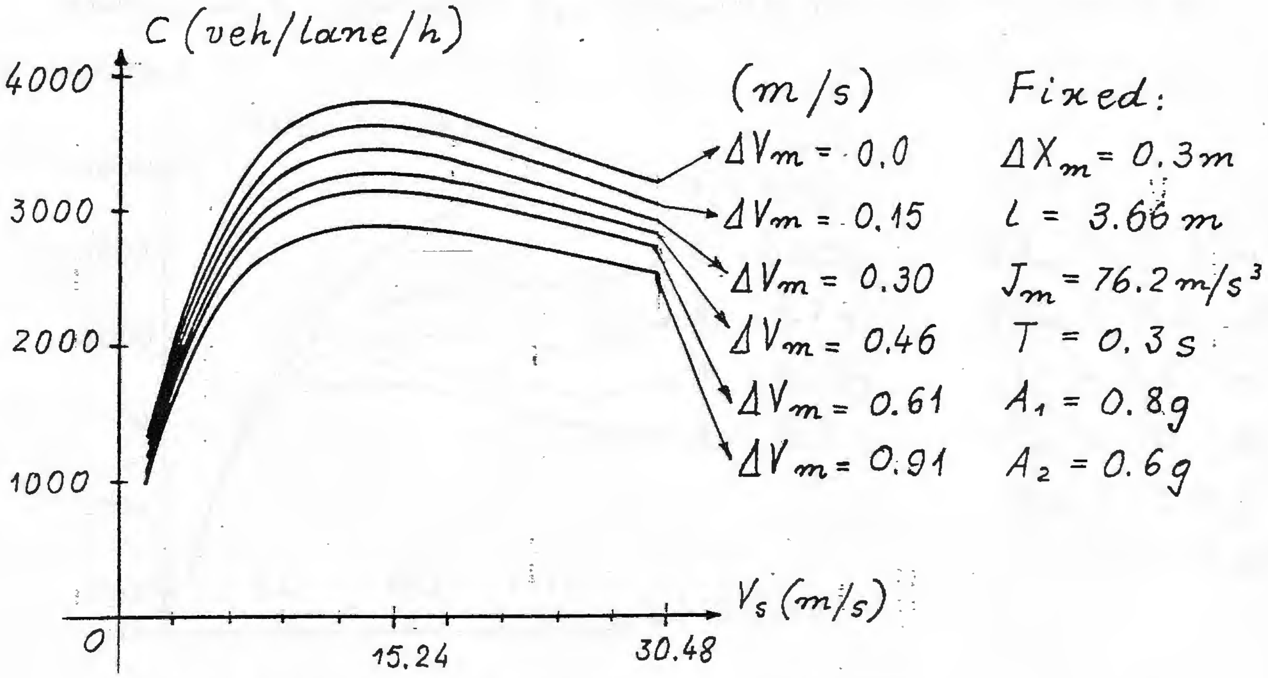


Figure 2.11 Maximum capacity C versus speed V_s with ΔV_m as a parameter.

The effects of varying A_2 with A_1 fixed ($A_1 = 0.8g$) are shown in Figure 2.12. It is clear that if A_2 is markedly smaller than A_1, the desired capacity is not safely achievable.

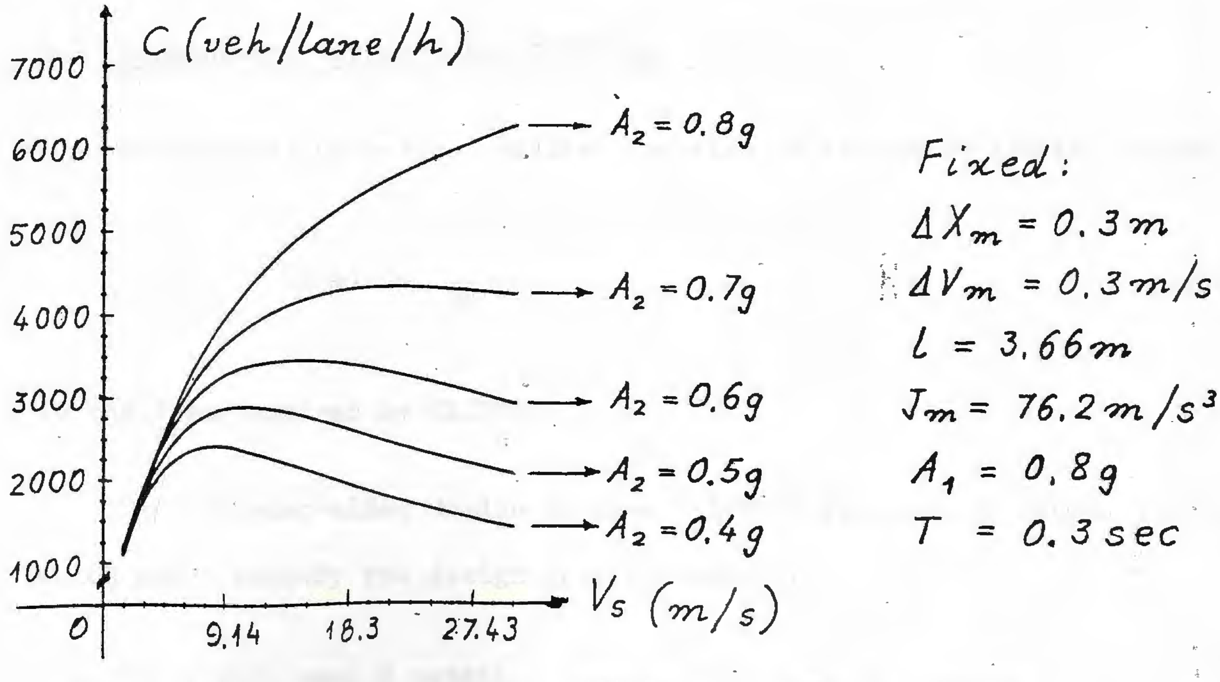


Figure 2.12 Maximum capacity versus speed with A_1 fixed and A_2 variable.

We can have a similar plot for A_2 fixed (0.6g) and A_1 variable, as shown in Figure 2.13 . For A_1 markedly larger than A_2 (e.g. $A_1 = 0.8g$, $A_2 = 0.6g$), the desired minimum capacity is not achieved at any speed. However, as A_1 approaches A_2, substantial increases in capacity are obtained.

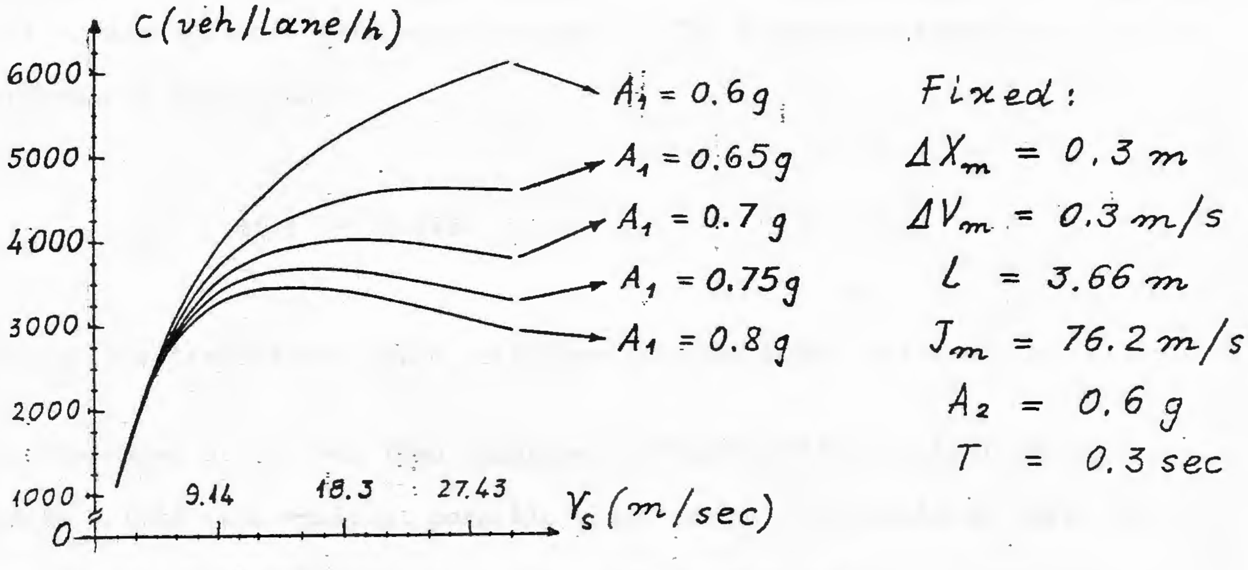


Figure 2.13 Maximum capacity versus speed with A_2 fixed and A_1 variable.

If, of course, lead-car deceleration were to exceed A_1 (i.e. $a_1(t) > A_1$) then condition (2.46) would be violated and collisions could result. The larger the choice of A_1, the less likely this will hapen. The same with A_2.

2.9 Compensator Design Using CRITERIA

The overall open-loop transfer function of the car-following system is:

$$G(s) = \frac{1}{4s + 1} \tag{2.51}$$

in the form required by CRITERIA.

The computer-aided design package CRITERIA was used to obtain a design which would satisfy the design criteria specified.

There were used 3 cases:

1. First a proportional and integral controller was employed of the form:

$$K(s,p) = p_1 + \frac{p_2}{s} \tag{2.52}$$

with the values of $D = 0.085045$ and $\varepsilon = 0.3048$. The initial parameters $p_1 = 1$ and $p_2 = 1$ were unconstrained. The system was stable and gave a performance functional

$$\phi(p) = 0.1789$$

without any iterations, which satisfied the inequality $\phi(p) < \varepsilon$.

2. The value of ε was then tightened to half of its original value, namely 0.1524 (the smallest possible value of ε that could be taken as a worst error). A P-I controller was used, again with initial values of $p_1 = 1$ and $p_2 = 1$. After two iterations the inequality was solved and gave final parameters

$$p_1 = 1.54 \quad \text{and} \quad p_2 = 1.1 \tag{2.54}$$

and a performance functional

$$\phi(p) = 0.1472 \tag{2.55}$$

3. Finally, a slightly more complicated compensator was tried of the form

$$K(s,p) = \frac{p_1(1+p_2 s)}{s(1+p_3 s)} \tag{2.56}$$

with initial values

$$p_1 = 1.4, \quad p_2 = 1.4, \quad p_3 = 0.1$$

The initial system was found to be stable and gave an excellent performance functional

$$(p) = 0.1393 \qquad (2.57)$$

without any iterations.

The plots of the system unit step response, the controller unit step response and the absolute error unit step response, for the three cases are shown in Figures 2.15, 2.16 and 2.17, and the numerical results in Figure 2.14 .

Results

Case A

NO	Functional	Value	Lower bound	Upper bound
1	P(1)	1.0000	-0.1000E+36	0.1000E+36
2	P(2)	1.0000	-0.1000E+36	0.1000E+36
1	Phi (1)	0.1789	0.0000E+00	0.3048
2	Phi (2)	5.2458	0.0000E+00	10.00
3	Abscissa of stab.	-0.2500	-0.1000E+36	-0.1000E-03

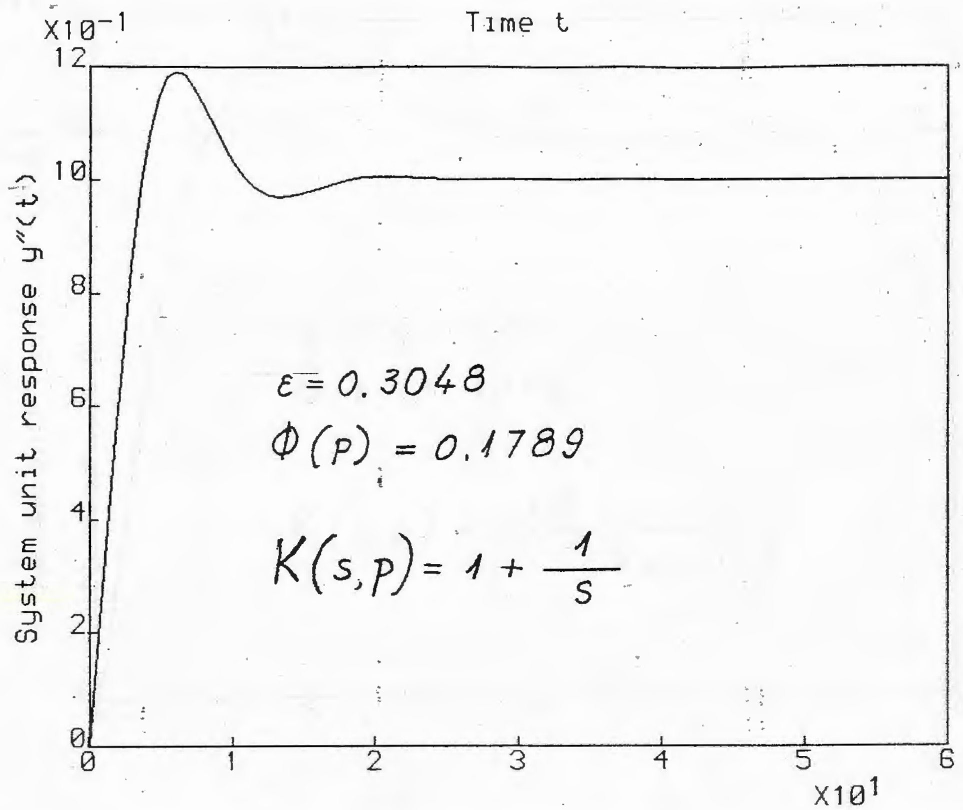
Case B

NO	Functional	Value	Lower bound	Upper bound
1	P(1)	1.5400	-0.1000E+36	0.1000E+36
2	P(2)	1.1000	-0.1000E+36	0.1000E+36
1	Phi (1)	0.1472	0.0000E+00	0.1524
2	Phi (2)	1.9393	0.0000E+00	10.00
3	Abscissa of stab.	-0.3367	-0.1000E+36	-0.1000E-03

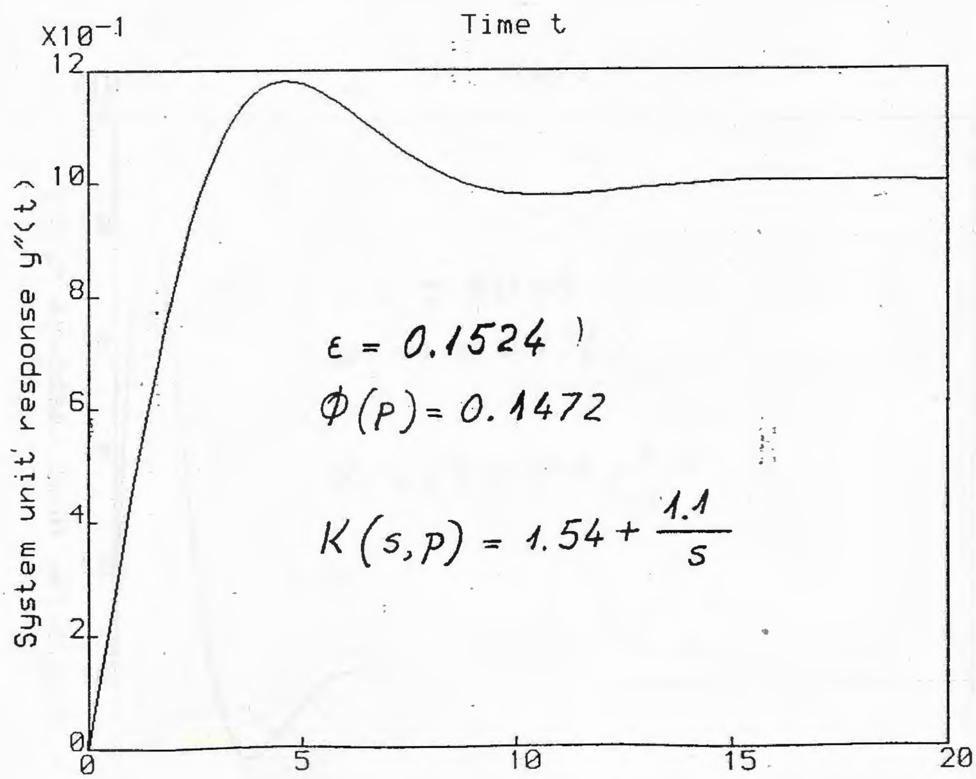
Case C

NO	Functional	Value	Lower bound	Upper bound
1	P(1)	1.4000	-0.1000E+36	0.1000E+36
2	P(2)	1.4000	-0.1000E+36	0.1000E+36
3	P(3)	0.1000	-0.1000E+36	0.1000E+36
1	Phi (1)	0.1393	0.0000E+00	0.1524
2	Phi (2)	1.9439	0.0000E+00	10.00
3	Abscissa of stab.	-0.3697	-0.1000E+36	-0.1000E-03

Figure 2.14 Numerical results of the three cases using CRITERIA.


X10^-1
Time t
12
10
8
6
4
2
0
System unit response y″(t)
0
1
2
3
4
5
6
X10^1
ε = 0.3048
Φ(p) = 0.1789
K(s,p) = 1 + 1/s

Case A


X10^-1
Time t
12
10
8
6
4
2
0
System unit response y″(t)
0
5
10
15
20
ε = 0.1524
Φ(p) = 0.1472
K(s,p) = 1.54 + 1.1/s

Case B

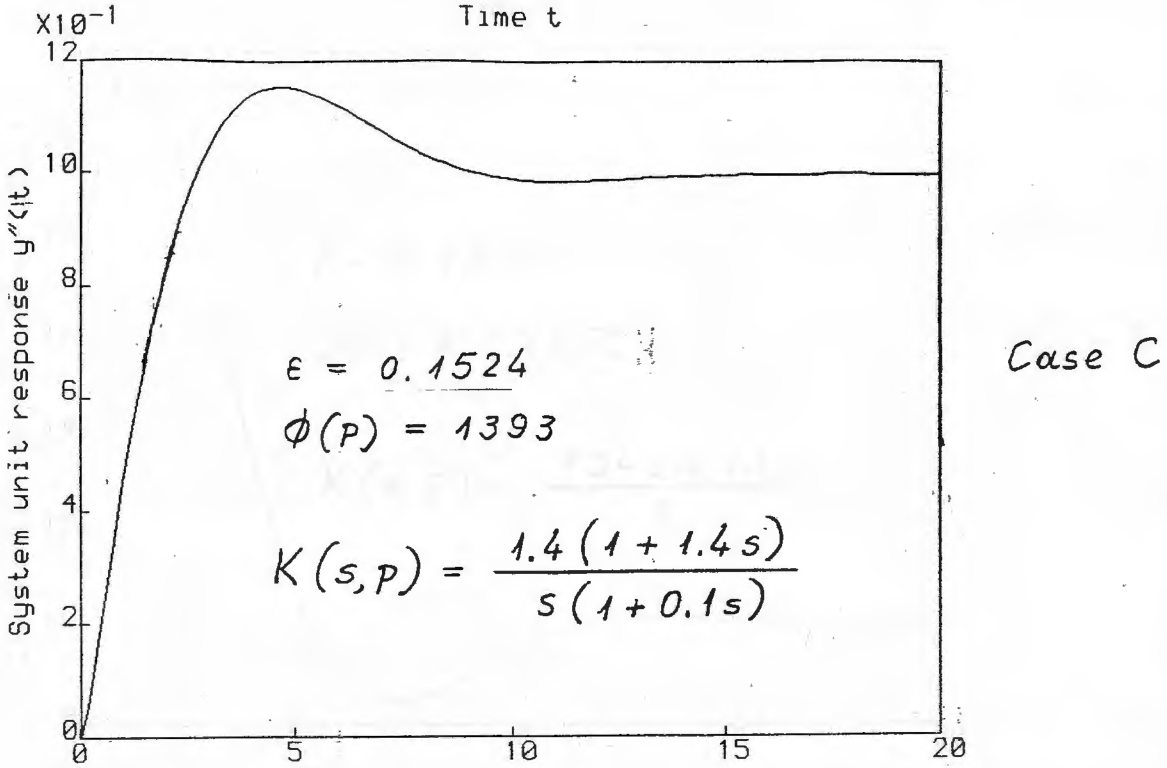
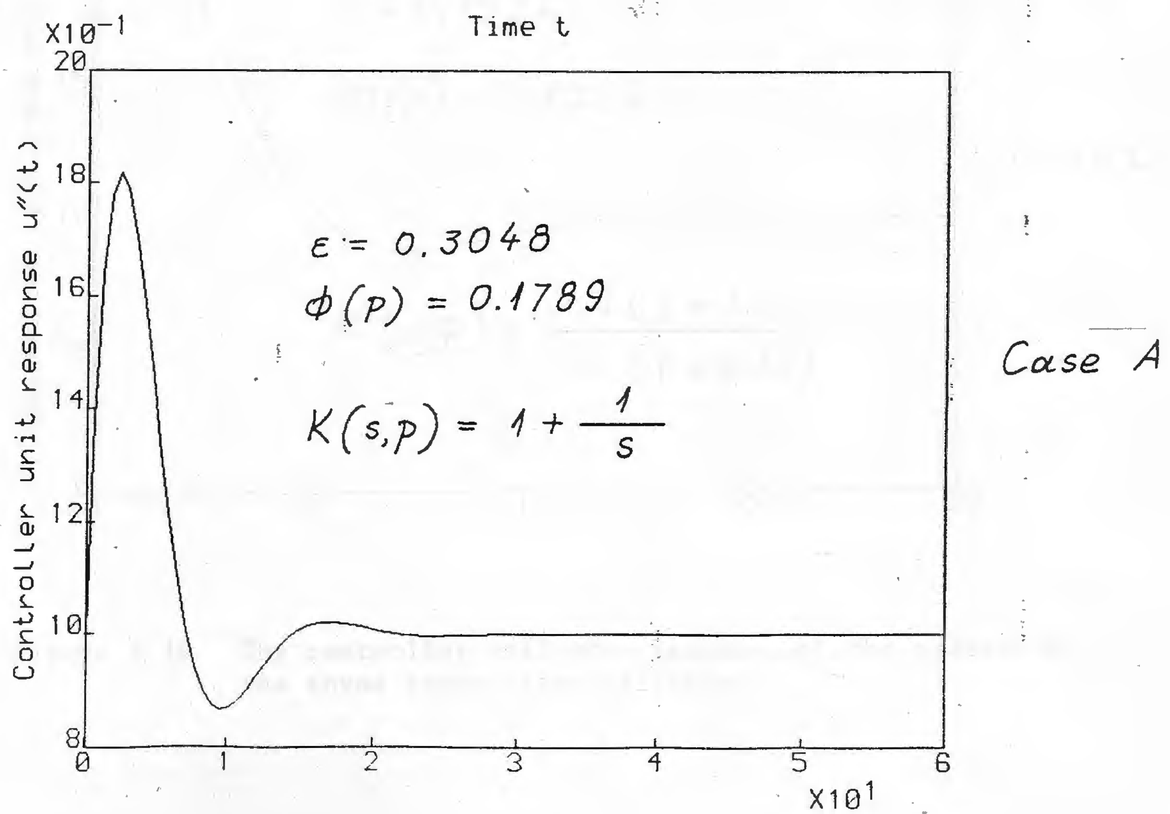
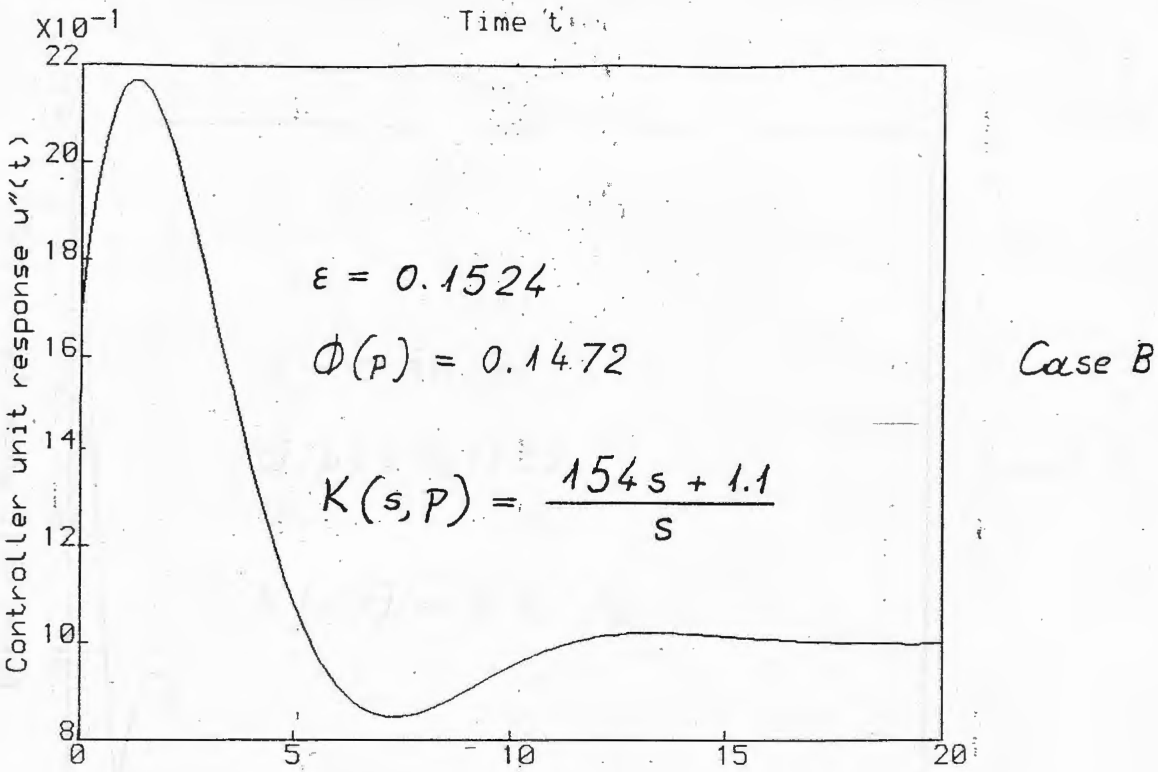


Figure 2.15 The system unit step response of the system in the three cases using CRITERIA





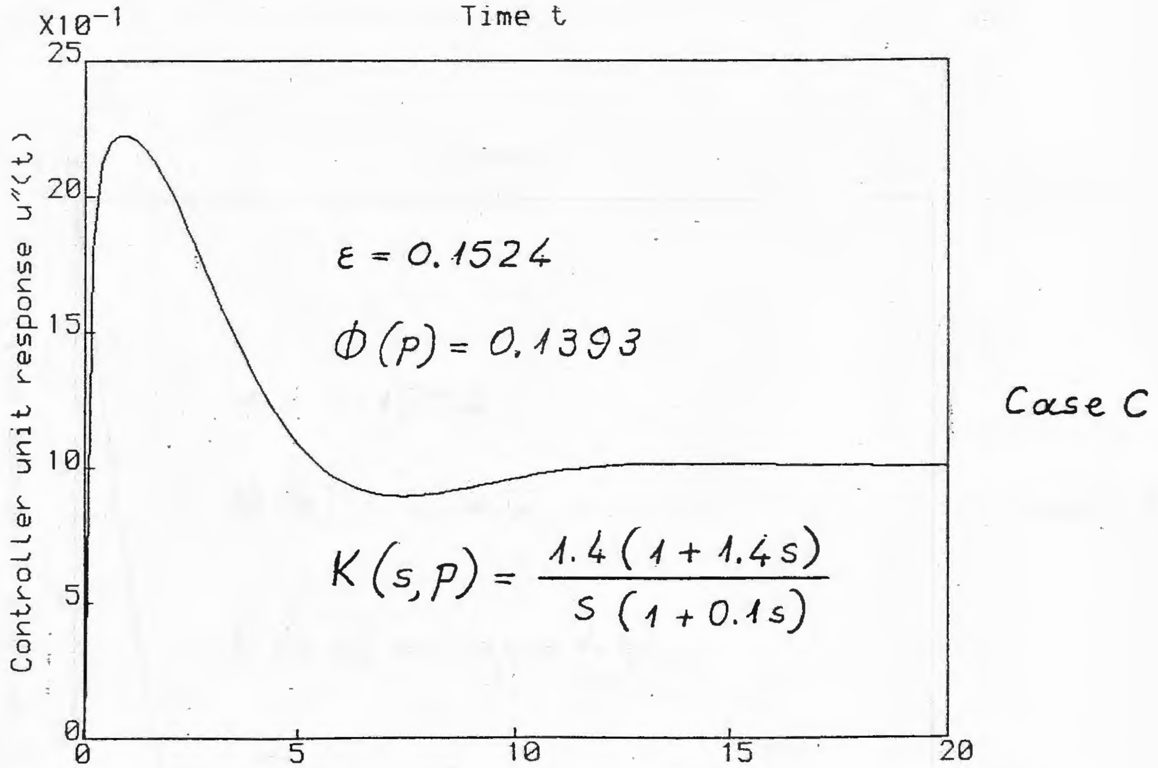
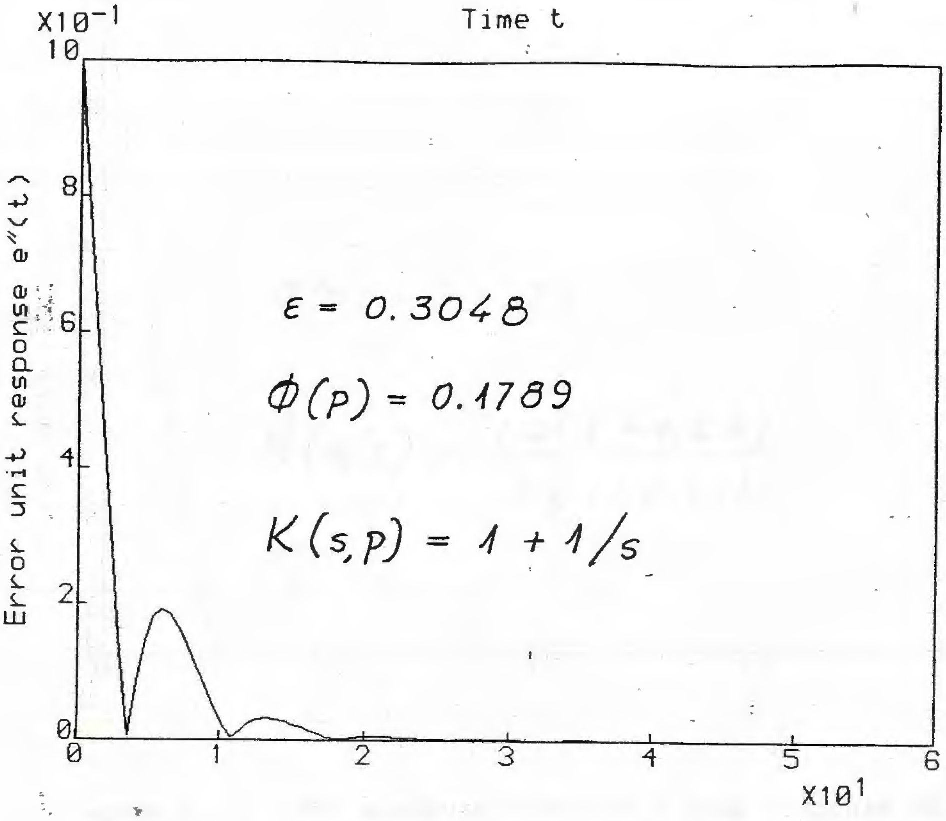
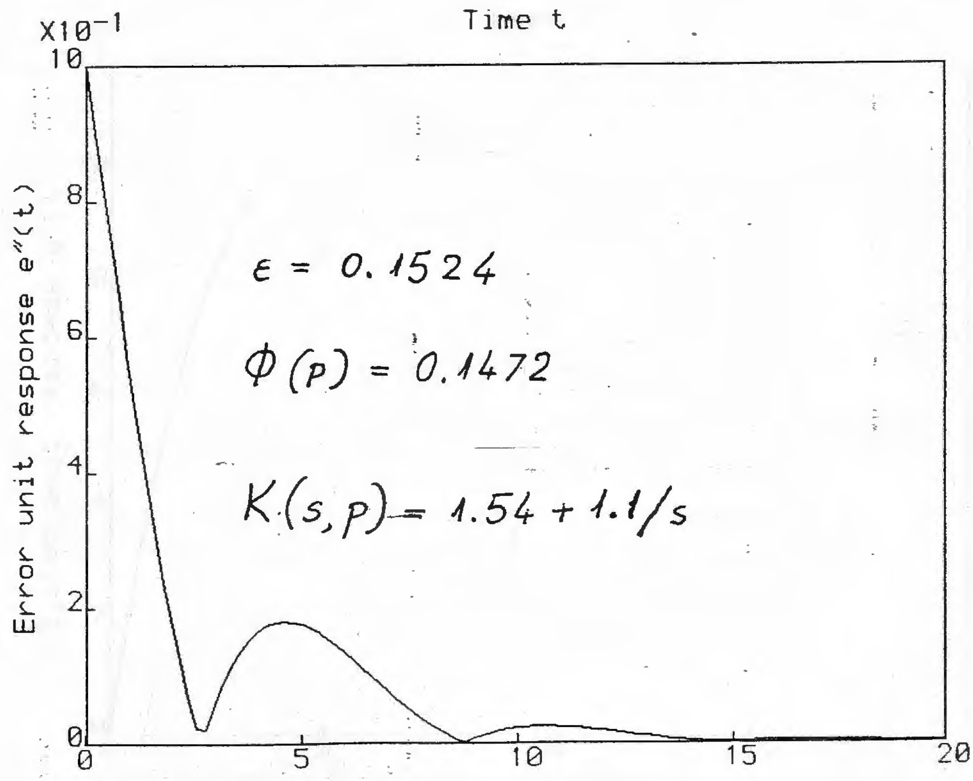


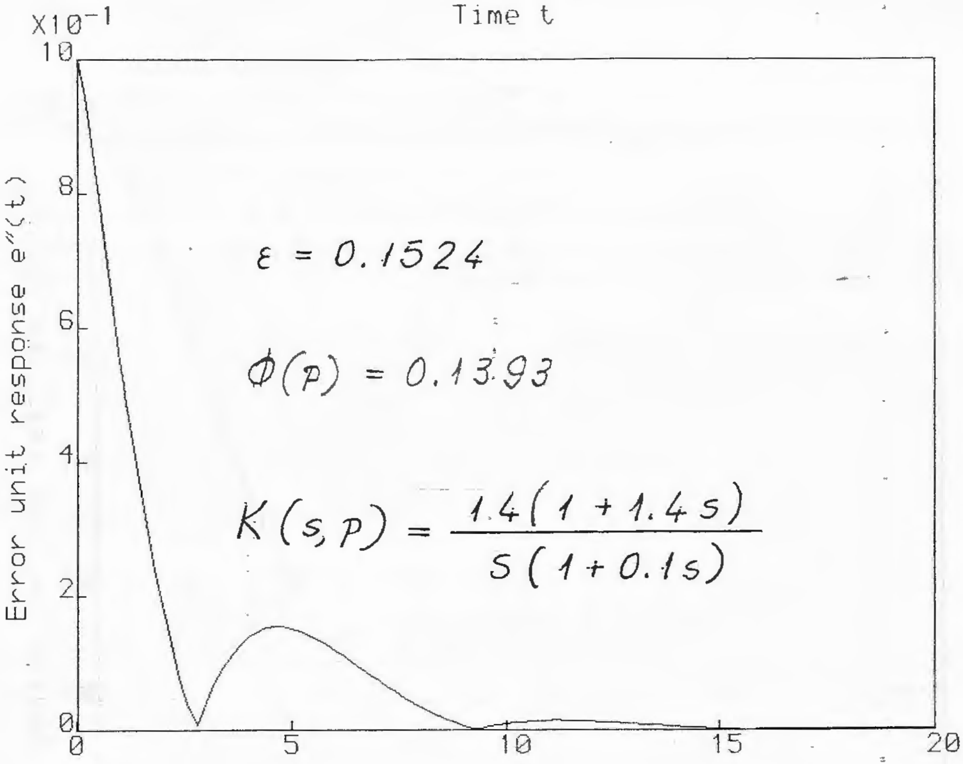
Figure 2.16 The controller unit step response of the system in the three cases using CRITERIA.


X10^-1
Time t
Error unit response e"(t)
10
8
6
4
2
0
0
1
2
3
4
5
6
X10^1
ε = 0.3048
Φ(p) = 0.1789
K(s,p) = 1 + 1/s

Case A


X10^-1
Time t
Error unit response e"(t)
10
8
6
4
2
0
0
5
10
15
20
ε = 0.1524
Φ(p) = 0.1472
K(s,p) = 1.54 + 1.1/s

Case B



Case C

Figure 2.17 The absolute error unit step response of the system in the three cases using CRITERIA.

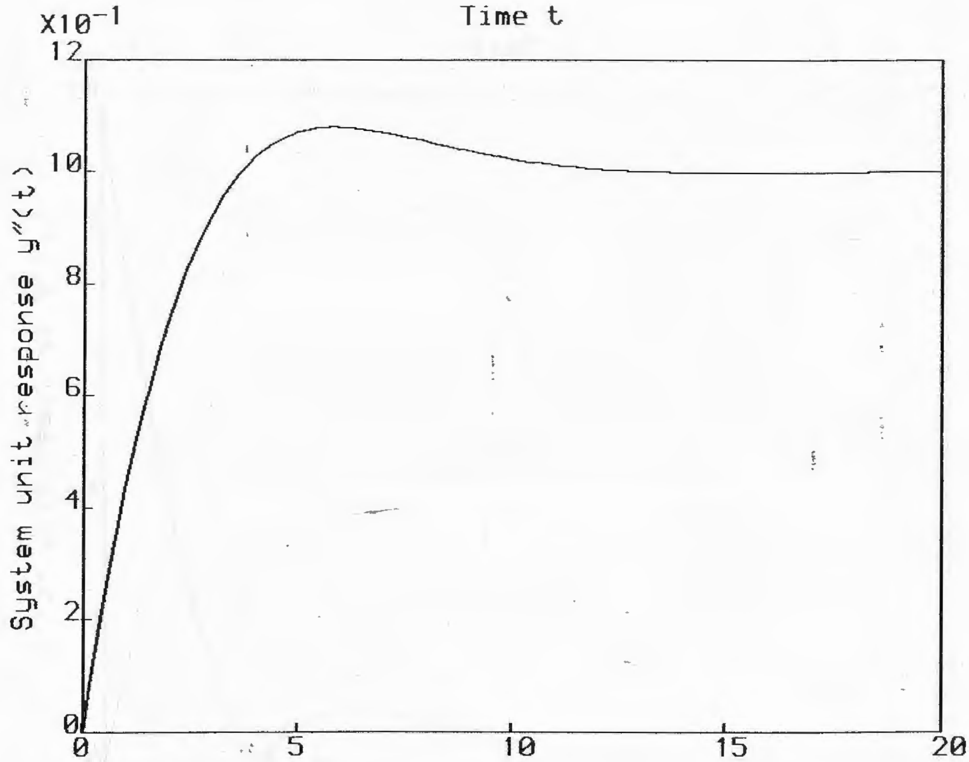


Figure 2.18 The system unit response of the system in the case study (Section 2.10).

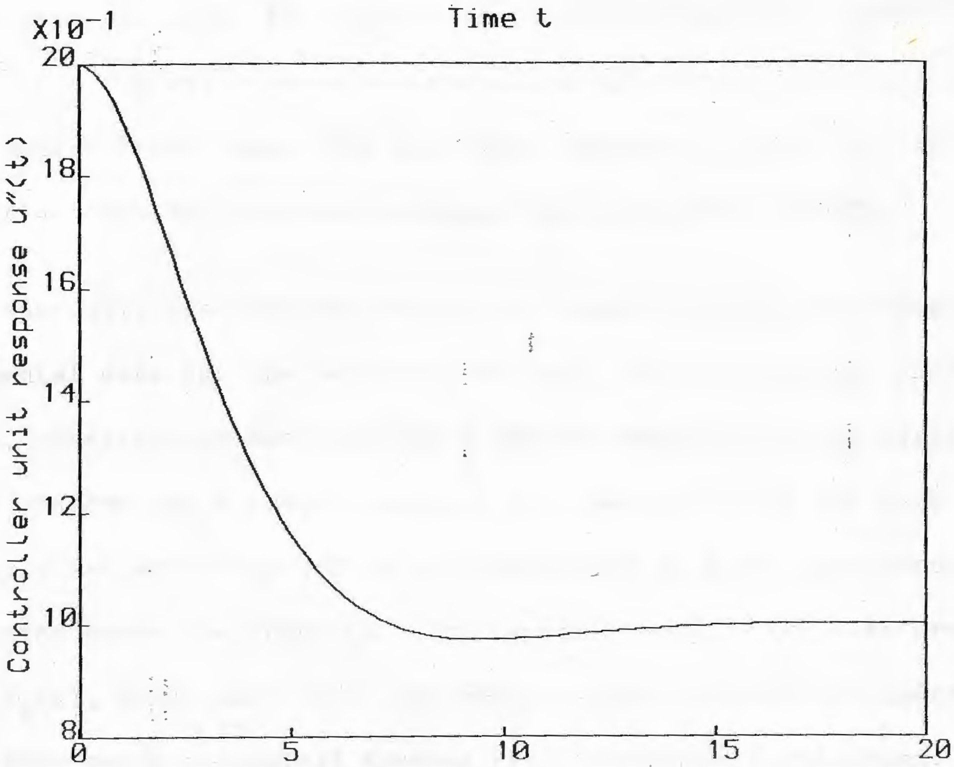


Figure 2.19 The controller unit step response of the system in the case study (Section 2.10).

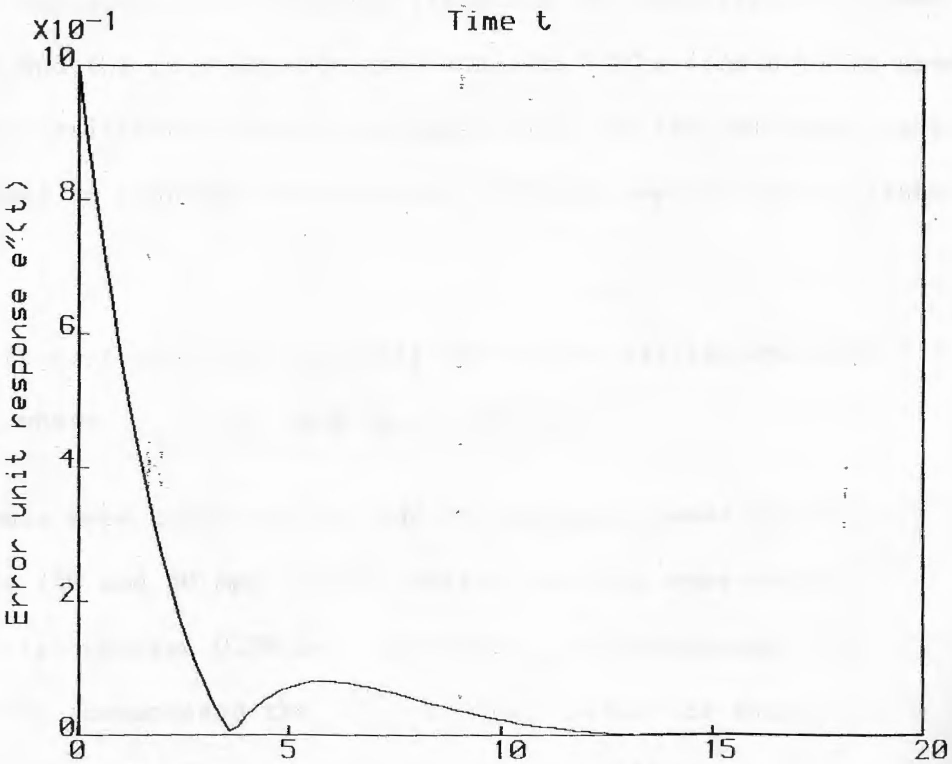


Figure 2.20 The absolute error unit step response of the system in the case study (Section 2.10).

2.10 Controller Designed in Case Study

In the case study the experimental work was done at a 5-mile unopened section of a typical highway which was a nearly ideal road with straight and approximately level lane. The test was conducted on days when the concrete road surface was dry and the wind speed was less than 16 km/h.

A specially instrumented vehicle was used to obtain the necessary experimental data for the velocity control. The main control functions - braking, acceleration and steering - were accomplished using electro-hydraulic control systems and a hybrid computer was installed over the back seat. All data collected were recorded on a 4-channel FM magnetic tape recorder, which was mounted under the computer. The required vehicle state information $v_1(t)$, $v_2(t)$, $h(t)$ and $v(t)$ was obtained via a tachometer geared to the drive-train and a mechanical take-up real/tachometer combination, which was attached between the controlled instrumented vehicle and a lead vehicle of the same make.

An automatic car-following situation was established between the lead vehicle and the instrumented test vehicle. The lead vehicle speed consisted of small variations about an average value and the following vehicle was programmed to respond in accordance with the overall system transfer function (2.17).

A proportional plus integral controller was implemented of the form (2.52), where $p_1 = 1.0$ and $p_2 = 2.0$.

Tests were conducted at lead car average speeds of 17.85 m/s and 26.8 m/s (40 and 60 mph respectively), and data were collected at various frequencies between 0.06 and 1.0 rad/s. This frequency band was chosen because it encompassed the critical frequencies for steady-state longitudinal control. The results are shown in Figure 2.21 .

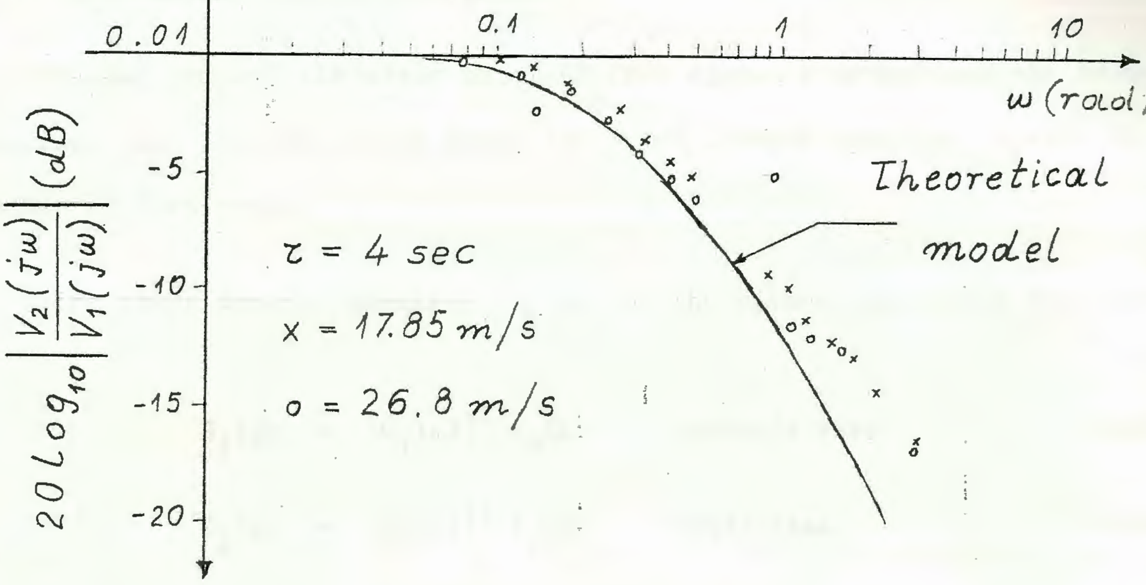


Figure 2.21 Linear velocity controller (Bode plot)

o - data obtained at average speed of 60 mph
x - data obtained at average speed of 40 mph.

This experiment was subsequently repeated four times on different roads in different ways, and under a variety of wind conditions ranging from both headwinds and crosswinds of at least ±4.6 m/s (16 km/h). In all cases, virtually the same results were obtained, which indicates the efficacy of the system compensation in overcoming environmental changes. There is also a close correspondence of the data at 17.85 and 26.8 m/sec which is a strong indication that the small-signal model is valid at other average speeds within this range.

Using CRITERIA for the transfer function (2.17) and the P-I controller 1.0+2.0/s, the plots obtained are shown in Figures 2.18, 2.19, 2.20 and the results in Figure 2.22.

NO	Functional	Value	Lower bound	Upper bound
1	P(1)	1.0000	-0.1000E+36	0.1000E+36
2	P(2)	2.0000	-0.1000E+36	0.1000E+36
1	Steady state error	0.0000	-0.1000E+36	-0.1000E-03
2	Rise time	2.8307	0.0000E+00	0.0000E+00
3	Overshoot (%)	0.0791	0.0000E+00	0.0000E+00
4	Settling time	10.1553	0.0000E+00	0.0000E+00
5	Undershoot (%)	0.0000	0.0000E+00	0.0000E+00
6	Min. contr. output	0.9718	0.0000E+00	0.0000E+00
7	Max. contr. output	2.0000	0.0000E+00	0.0000E+00
8	Abscissa of stabi.	-0.3750	-0.1000E+36	-0.1000E-03

Figure 2.22 Numerical results of the case study car-following system.

2.11 Spectrum of Output in Both Cases

We can compare the error obtained from Zakian's method and the error obtained from the case study using the power density spectrum $S_y(\omega)$ of the output in both cases.

The power density spectrum $S_y(\omega)$ of the output, and hence the error, is

$$S_1(\omega) = |G_1(\omega)|^2 S_w(\omega) \qquad \text{Zakian's case} \tag{2.58}$$

$$S_2(\omega) = |G_2(\omega)|^2 S_w(\omega) \qquad \text{Study case} \tag{2.59}$$

where $G(s)$ is given by (2.51),

$$G_1(s) = \frac{G(s)}{1+G(s)K_z(s,p)} \qquad s = j\omega \tag{2.60}$$

$$G_2(s) = \frac{G(s)}{1+G(s)K_s(s,p)} \qquad s = j\omega \tag{2.61}$$

$$K_z(s,p) = 1.54 + 1.1/s \ ,$$

$$K_s(s,p) = 1 + 2/s$$

and

$$S_w(\omega) = 0.1063/(1+4\omega^2) \tag{2.62}$$

Thus,

$$G_1(s) = \frac{s}{4s^2+2.54s+1.1} \tag{2.63}$$

Similarly,

$$G_2(s) = \frac{s}{4s^2+2s+2} \tag{2.64}$$

Hence, substituting (2.62), (2.63) and (2.64) into (2.58), (2.59),

$$S_1(\omega) = \left|\frac{j\omega}{4(j\omega)^2+2.54(j\omega)+1.1}\right|^2 \cdot \frac{0.1063}{1+4\omega^2}$$

$$= \frac{0.1063}{\{6.4516\omega^2+(1.1-4\omega^2)^2\}(1+4\omega^2)} \tag{2.65}$$

and similarly,

$$S_2(\omega) = \frac{0.1063}{\{4\omega^2+(2-4\omega^2)^2\}(1+4\omega^2)} \tag{2.66}$$

The variance σ_i^2 of the output and thus the error, is:

$$\sigma_i^2 = \frac{1}{\pi}\int_0^\infty S_i(\omega) \qquad i = 1,2 \tag{2.67}$$

(1 represents the Zakian case, and 2 the study one)

Integrating numerically, using the Simpson rule after substituting (2.65) into (2.67), we get:

$$\sigma_1^2 = 6.4059\text{x}10^{-4} \tag{2.68}$$

and the standard deviation is

$$\sigma_1 = 0.02531 \tag{2.69}$$

This standard deviation gives a maximum error of $3\sigma_1$,

$$3\sigma_1 = 0.07593 \tag{2.70}$$

Following the same procedure for σ_2 we get, for the case study the maximum error $3\sigma_2$,

$$3\sigma_2 = 0.102424 \tag{2.71}$$

Thus the design achieved by using Zakian's method is an improvement on that achieved in the case study.

We could calculate the error functional $\phi(p)$ for the input space bounded by $D = 0.085045$ by hand-computation using the fundamental theorem:

$$\phi(p) \quad = \quad ||e(h,p)||_1 . D \qquad (2.72)$$

where
$$||e(h,p)||_1 \quad = \quad \int_0^\infty |e(t,h,p)|dt \qquad (2.73)$$

and $e(t,h,p)$ has the Laplace transform:

$$E(s,\delta,p) \quad = \quad \frac{1}{1+G(s)K(s,p)} \cdot \frac{1}{s} \qquad (2.74)$$

Equation (2.74) for $G(s) = 1/4s+1$ and $K(s,p) = \dfrac{1.54s+1.1}{s}$ becomes:

$$E(s,\delta,p) \quad = \quad \frac{4s+1}{4s^2+2.54s+1.1} = \frac{s+0.25}{s^2+0.635s+0.275} \qquad (2.75)$$

or after some manipulation in the denominator:

$$E(s,\delta,p) \quad = \quad \frac{s+0.3175}{(s+0.3175)^2+0.4174^2} - 0.1617 \frac{0.4174}{(s+0.3175)^2+0.4174^2} \qquad (2.76)$$

Taking inverse Laplace transforms, equation (2.76) becomes:

$$e(t,h,p) \quad = \quad e^{-0.3175t}[\cos(0.4174t)-0.1617\sin(0.4174t)] \qquad (2.77)$$

where 0.4174 is expressed in rad/s.

Substituting (2.77) into (2.73) and integrating numerically, using Simpson's rule, we find:

$$||e(h,p)||_1 \quad = \quad 0.90827 \qquad (2.78)$$

so that (2.72) becomes:

$$\phi(p) \;=\; 0.077244 \tag{2.79}$$

which is just a little larger than the maximum error obtained from (2.70). This is because, as mentioned before, using power spectra, we obtain an input space which is a subspace of the original one, so the error functional will be smaller than the real one. (The result of (2.79) is not exactly the real one because it contains D which is not the real value obtained through power spectrum of stationary processes.)

DISCUSSION AND CONCLUSION

A velocity controller was designed for an automatic car-following system in a steady-state mode using Zakian's method, and a comparison was made between this method and the case study one. In this system both methods gave an appropriate design. However, the advantage of Zakian's approach is that a precise measure of control is obtained.

In the case study there were tried five other control models of the general form (2.18), and the controller used for the case experiment had a reasonable time constant $\tau = 4$ sec. However, if the use of headway feedback is contemplated, one must be quite careful in the choice of a linear controller, if one wishes to utilize such feedback and also achieve high traffic densities $(k = k_3+k_4 \leq 1)$. In this case he must use a vehicle system with a small time constant τ so that the system can respond very quickly to small command changes in speed. One of the five other controllers used in this study was that, where $k_1 = 1$, $k_2 = 0.5$, $k_3 = 0$ and $k_4 = 1$. With these values of k_i the transfer function (2.20) becomes:

$$G(s) = \frac{1}{s+1}$$

and the corresponding values of τ and k are both 1. This choice of parameters resulted in both asymptotic stability and possible high traffic flow rates, but it also led to some difficulties, i.e. the controlled vehicle would be highly responsive to changes in lead-vehicle speed, resulting in passenger discomfort and poor gas mileage. So, in order to ensure system compatibility with existing traffic, the vehicle system time constant τ should be chosen so that it may be realized by the majority of automobiles in use of the highways.

However, for high lane capacities ($\geq$ 3600 veh/lane/h) criteria specifications have been described in Section 2.7, where the spacing would correspond to a time headway of 1 sec or less for a uniform spacing headway policy. With this policy the minimum permitted intervehicular spacing is determined by safe controlled stopping distance considerations. These considerations are determined by vehicle characteristics (e.g. its control system, its braking system and the vehicle/roadway interface), as well as by network control factors.

Now, for the final design of the compensator, CRITERIA package was used, based on Zakian's framework. There were made three attempts, which gave three different values for the system performance.

First, a P-I controller was used with a good enough upper bound $\varepsilon = 0.3048$ specified in Section 2.8 . The system was found stable and gave a performance functional $\phi(p) = 0.1789$ without any iterations. Then the same kind of P-I controller was tried, but the upper bound was tightened to the minimum permissible value specified in 2.8, namely $\varepsilon = 0.1524$. By the method of inequalities, after only two iterations, the system was found stable and gave a performance functional $\phi(p) = 0.1472$ which satisfied the performance criterion. A third attempt was made by trying a more complex compensator of the form (2.56) . This gave a very good performance functional $\phi(p) = 0.1395$ without any iterations.

These three attempts were made just to show how easy it is, using CRITERIA, to get the most suitable parameters required for a good design of the compensator used to meet the performance criterion. In all these cases the upper bound $D = 0.085045$ of the absolute rate of change of inputs was used, estimated as three times the standard deviation of the rate of change of wind disturbance input as shown in Section 2.7 .

It must be pointed out here, that the major difficulty using Zakian's method is the determination of the input space, which is perhaps the most important part of the design, the importance depending mainly upon the criticality of the system.

Many systems are subjected to random disturbances, which may not have a finite bound on their derivative. The problem is to estimate a reasonable bound for D with a great amount of confidence. The statistics of extremes [19] give a method to do this, but in problems of natural disturbances, such as winds, waves, rains and so on, records of extreme samples need to be kept at least 20 years in order to be implemented with much confidence. The classical method of dealing with stochastic inputs is to use properties such as power density spectrum of stationary processes and the probability density function. Thus we need to make some assumptions about the nature of the disturbances in order to get a model for these disturbances, which can then be analysed. However, using power density spectra to determine D is not always the ideal method, because the assumption that the disturbance is Gaussian stationary, is not always the case; it can only be an approximation to what occurs in reality. In this case, as it has been emphasized in (1.7B) Zakian's input space F becomes in fact a subspace because of this restriction of stationary process. When the input space is the originally proposed one by Zakian without any assumptions, but the only one restriction that the absolute rate of change of the inputs must be bounded, which is in fact a physical law, only then can Zakian's theory be implemented, especially in the control of critical systems.

The resulting design from Zakian's method was an improvement on that of the case study and this because the design parameters for the actual compensator were worked out through the method of inequalities and the incorporated input space in CRITERIA software package, so that they gave the desired physical performance based on Zakian's criteria, whereas in the case study a trial was made with some specific parameters and the system performance was "measured" by using Bode plots. Moreover, the corresponding study was conducted in a variety of external windspeeds of up to ±4.5 m/s (16.2 km/h), which is obviously not the maximum wind speed that could happen.

Generally, traditional methods, such as root locus, Nyquist diagrams, Bode plots, etc., contend with difficulties for achieving the design of critical systems using various criteria, such as phase or gain margin, rise time, overshoot, settling time, damping ratio, etc., and the actual compensator design is often a tedious process, even with the aid of a computer. The criteria used are sometimes far from the real ones and the optimal solutions could be unnecessarily complex and, hence, expensive for implementation. In Zakian's case, once the input space design criteria and the system model have been determined, then the actual design is a simple process using the powerful CRITERIA package.

Thus, Zakian's design framework is well suited to the design of any control system, and especially of a critical one.

REFERENCES

[1] Astrom, K.J. "Introduction to stochastic control theory". Academic Press, 1970.

[2] Bada, A.T. "Design of delayed control systems with Zakian's method". Int. J. of Control, vol. 40, No. 4 (1984), pp. 773-782.

[3] Bada, A.T. "Criteria - a package for computer-aided control system design". Control Systems Centre Report No. 643, UMIST, Manchester, 1985.

[4] Bender, J.G. "Experimental studies in vehicle automatic control". Commun. and Contr. Syst. Lab. Eng. Exp. Station, Ohio State Univ. Columbus Rep. EES 276A-5, 1968.

[5] Bender, J.G., Fenton, R.E. "A study of automatic car-following". IEEE Trans. Veh. Technol., vol. VT-18 No. 3 (1969), pp. 234-140.

[6] Bender, J.G., Fenton, R.E. and Olson, K.W. "An experimental study of vehicle automatic longitudinal control". IEEE Trans. on Vehicular Technology, Vol. VT20, No. 4, 1971, pp. 114-123.

[7] Blackwell, L.M. "A study of the adaptability of the automobile to automatic control". Commun. and Contr. Syst. Lab. Ohio State Univ. EES 276A-3, 1967.

[8] Brebbia, C.A. et al. "Vibrations of engineering structures". Computational Mechanics Ltd., pp. 366-377.

[9] Chandler, R.E., Herman, R., Montroll, E.W. "Traffic dynamics : studies in car-following". Oper. Res., vol. 6, No. 2, (1958), pp. 165-184.

[10] Cooper, R. "Atmospheric turbulence with respect to moving ground vehicles". Journal of Wind Engineering and Industrial Aerodynamics, vol. 17 (1984), pp. 215-238.

[11] Fenton, R.E. "Automatic vehicle guidance and control - a state of the art survey". IEEE Trans. on Veh. Technol., vol. VT-19, No. 1, 1970, pp. 153-161.

[12] Fenton, R.E. "A headway safety policy for automated highway operations". IEEE Trans. on Veh. Technol., vol. VT-28, 1979, pp. 22-28.

[13] Fenton, R.E., Cosgriff, R.L., Olson, K.W. and Blackwell, L.M. "One approach to highway automation". IEEE Proceedings, vol. 56, No. 4, 1968, pp/ 556-566.

[14] Fenton, R.E. et al. "Fundamental studies in the automatic longitudinal control of vehicles". Dep. Elec. Eng. Ohio St. Univ. Columbus Rep. DOT-TST-76-79 (1975).

[15] Fenton, R.E., Olson, K.W., Bender, J.G. "Advances toward the automatic highway". Highway Res. Board, Washington, D.C., 1971, pp. 18-22.

[16] Fenton, R.E., Montano, W.B. "An intervehicular spacing display for improved car-following performance". IEEE Trans. Man-Machine Syst., vol. MMS-9, 1968, pp. 29-35.

[17] Glimm, J., Fenton, R.E. "An accident-severity analysis for a uniform-spacing headway policy". IEEE Trans. on Veh. Techn., vol. VT-29 No. 1, (1980), pp. 96-103.

[18] Grimmett, G., Stirzaker, D. "Probability and Random Processes". Oxford Science Publications 1985.

[19] Gumbel, E.J. "Statistics of extremes". Columbia University Press 1958.

[20] Hajdu, L.P., Gardiner, K.W., Tamura, H. and Pressman, G.L. "Design and control considerations for automated ground transportation systems". IEEE Proceedings, vol. 56, No. 4 (1968), pp. 493-513.

[21] Holmes, J.D. "Non-Gaussian characteristics of wind pressure fluctuations". Journal of Wind Engineering and Industrial Aerodynamics, vol. 7 (1981), pp. 103-108.

[22] Lenard, M. "Safety considerations for a high density automated vehicle system". Transport Sci., vol. 11, 1970, pp. 138-158.

[23] Lipschutz, S. "Probability". Schaum's Outline Series, 1974.

[24] Lukas, M.P. "Advances in instrumentation". Instrument Society of America Annual Conference, vol. 26, Part 1 (1971), Paper No. 71-508.

[25] Moriarty, W. "On the estimation of extreme wind gusts by direction sector". Journal of Wind Engineering and Industrial Aerodynamics, vol. 13 (1983), pp. 127-138.

[26] Newton, G.C., Gould, L.A., Kaiser, J.F. "Analytical design of linear feedback controllers". John Wiley & Sons, 1957.

[27] Peppard, L.E., Gourishankar, V. "An optimal automatic car-following system". IEEE Trans. on Veh. Technol., vol. VT-21, No. 2, 1972, pp. 67-73.

[28] Pipes, L.A. "An operational analysis of traffic dynamics". J. Appl. Phys., vol. 24 (1953), pp. 274-281.

[29] Reed, D., Scanlan, R. "Autoregressive representation of longitudinal, lateral and vertical turbulence spectra". Journal of Wind Engineering and Industrial Aerodynamics, vol. 17 (1984), pp. 199-214.

[30] Rose, M.J. "Commercial vehicle fuel economy - the correlation between aerodynamic drag and fuel consumption of a typical truck". Journal of Wind Engineering and Industrial Aerodynamics, vol. 9 (1981), pp. 89-100.

[31] Roussillon, G. "Contribution to accurate measurement of aerodynamic drag on a moving vehicle from coastdown tests and determination of actual rolling resistance". Journal of Wind Engineering and Industrial Aerodynamics, vol. 9, (1981), pp. 33-48.

[32] Rule, R.G., Fenton, R.E., Bender, J.G. "On modelling vehicle dynamics - a time series analysis approach". Advance in Instrumentation, vol. 26, Part 1, ISA 1971, pp. 71-509.

[33] Sachs, P. "Wind forces in engineering". Pergammon Press, 1972.

[34] Spiegel, M.R. "Probability and Statistics". Schaum's Outline Series, 1980.

[35] Telionis, D., Fahrner, C. and Jones, G. "An experimental study of highway aerodynamics interferences". Journal of Wind Engineering and Industrial Aerodynamics, vol. 17 (1984), pp. 267-293.

[36] Todosiev, E.P. "An automobile simulator". AIEE Empire Tri-District Meeting, Erie, Pa., 1962, Paper No. DP62967.

[37] Wilson, D.G. "Automated guideway transportation between and within cities". M.I.T. Rep. FRA-RT-72-14, 1971, (PB 206-269).

[38] Zakian, V., Al-Naib, U. "Design of dynamical and control systems by the method of inequalities". IEE Proceedings, vol. 120, No. 11, (1973), pp. 1421-1427.

[39] Zakian, V. "Computation of abscissa of stability by repeated use of the Routh test". Control Systems Centre Report No. 441, UMIST, Manchester 1978.

[40] Zakian, V. "New formulation for the method of inequalities". IEE Proceedings, vol. 126, No. 6, (1979), pp. 579-584.

[41] Zakian, V. "A criterion of approximation for the method of inequalities" Int. J. of Control, vol. 37, No. 5 (1983), pp. 1103-1112.

[42] Zakian, V. "A performance criterion". Int. J. Control, 1986, vol. 43, No. 3, pp. 921-931.

[43] Zakian, V. "On performance criteria". Int. J. Control, 1986, vol. 43, No. 4, pp. 1089-1092.

[44] Zakian, V. "Input spaces and output performance". Report No. 650, Control Systems Centre, UMIST, Manchester, 1986.

[45] Zakian, V. "Design formulations". Control Systems Centre Report No. 651 UMIST, Manchester, 1986.

ACKNOWLEDGMENTS

The author wishes to thank his supervisor Dr. V. Zakian, for his inspiring comments and helpful discussions; Mr. A. Jaimoukha for his useful suggestions concerning this work, as well as Mrs. M.V. Butterworth for typing the manuscript.