Vehicle Scheduling and Control in Personal Rapid Transit Systems

3 reports

Report of Project AERO-1
Submitted to University Grant Division
Urban Mass Transportation Administration
U.S. Department of Transportation
Under Grant MN-11-0003

Program in Urban Transportation
Center for Urban and Regional Affairs
University of Minnesota
September, 1973

This report covers three studies of various aspects of the scheduling and control of PRT vehicles!

The Simulation of a PRT System Operating Under Quasi-Synchronous Control

by

Harold L. York*

Introduction

The routing and scheduling of vehicles is an important problem in any demand-activated transit system. Vehicles should be dispatched and routed to their destinations over paths which minimize travel time but do not create congestion problems in the network. Empty vehicles must be shuttled in a manner which provides a vehicle within a reasonable time after a demand for service has been made, and the movements of these empty vehicles should be optimized to reduce fleet requirements and operating costs.

Some of these operational problems can be studied successfully with analytical methods. Using stochastic models, useful information can be obtained concerning the operation of an individual station or interchange.(1), (2) However, the performance of an operational strategy in a network composed of many stations and interchanges, in which demands for service occur in a random manner, can only be studied adequately by computer simulation.

Several operational strategies for PRT systems have been proposed. These schemes, especially when the routing of vehicles is considered, depend on whether the system is operating under synchronous, quasi-synchronous, or asynchronous control. Early studies indicated that systems operating under fully-synchronous control would have low capacity under most circumstances, and that computer logic and memory requirements would be extensive. More recently, several investigators have shown that the synchronous system can be modified to significantly improve performance (3), (4). In this modification, groups of moving slots are combined into units called cycles. Vehicle maneuvers within a cycle are permitted. This relaxation of the fully-synchronous concept significantly increases the capacity of the system.

Under quasi-synchronous or asynchronous control, the vehicle path through the network is not precisely predetermined. Future switching states are not reserved for the vehicle, and the possibility of merging conflicts at interchanges exists. These conflicts are normally resolved by permitting the vehicles at or near the interchange to maneuver for an open position on the main line. Occasionally, it may be necessary to deny a

*Assistant Professor, Department of Aerospace Engineering and Mechanics, University of Minnesota

vehicle access to the turn ramp. In this case, the vehicle must be rerouted to its destination. The probability of a vehicle being rerouted at any interchange in the network must be very small if the system is to perform well. Operational strategies for systems under quasi-synchronous control have been developed and simulated by several research groups (5),(6).

System Description

The purpose of this paper is to describe an operational strategy for a PRT system operating under quasi-synchronous control, and to demonstrate by the use of computer simulation the feasibility and limitations of this strategy in a very high-demand situation. The computer simulation was developed to model the detailed operation of a small-vehicle transit network, and to determine the effects of random passenger demands on system performance. The simulation includes a dynamic minimum-path routing scheme and an optimal procedure for shuttling empty vehicles in the system. Slot slipping is performed on interchange and station ramps. With minor exceptions, all other links in the system are assumed to operate in a synchronous manner.

The network is modeled by dividing each section of the guideway into slots of one headway length. Every slot on the main network lines as well as those on the station and interchange ramps is assigned a storage location in the computer, and each station in the system is assigned a number. Each storage location which represents a guideway slot will contain the destination-station number of the vehicle which occupies the corresponding slot in the network at that time. If the slot is unoccupied, a zero is assigned to the location. At time increments of one headway time, the whole system is updated. The updating involves shifting or otherwise modifying the contents of the storage locations representing the guideway slots to simulate the movements of vehicles through the system. At each time-step, demands for vehicles at stations are also generated using a Monte Carlo method. In the network example described later, a headway and update time of one second is used.

The updating of vehicle positions on the main lines of the network is performed by the following sequence of logic operations:

(1) If the vehicle is in a slot which does not contain a demerge switch, it is advanced one slot on the main line.

(2) If an interchange demerge switch is occupied, the vehicle's destination is interrogated. The destination number is then checked with a minimum-path table to determine whether the vehicle should proceed on the main line or be placed on the interchange ramp.

(3) If the vehicle is located in a station demerge switch slot, its destination is interrogated. If that destination corresponds to the appropriate station number, the vehicle is

placed on the station entrance ramp. Otherwise, the vehicle is moved one slot forward on the main line.

In a quasi-synchronous control scheme, vehicle maneuvering takes place on station and interchange ramps. Interchange ramps may be divided into three sections consisting of a deceleration ramp, a maneuvering section, and an acceleration ramp. On both the acceleration and deceleration ramps, the vehicle operates in a synchronous manner. In the maneuvering section, the vehicle may adhere to a constant-speed profile, which is normally lower than the main-line speed, or it may "slip" slots by a combination of acceleration and deceleration maneuvers. The logic operations by which the interchanges are updated at each time step can be enumerated as follows:

(1) If the lead slot of the maneuvering section is unoccupied, all vehicles in this section are advanced one slot.
(2) If the lead slot is occupied and an open slot on the main line is available for this vehicle, the vehicle is placed in the first slot of the synchronous acceleration ramp, and all vehicles behind it are advanced one slot.
(3) If the lead slot is occupied but an open slot on the main-line is not available, the vehicle must slip a slot, and all vehicles behind the lead vehicle, up to the first vacant slot, must also slip a slot.
(4) A vehicle is permitted onto the interchange ramp if there is room for the vehicle in the maneuvering section.

In the simulation, only a specified number of vehicles are allowed to slip slots simultaneously on the maneuvering section. In the example described later, up to ten vehicles are permitted to slip slots at any time. There will be space for a vehicle on the maneuvering section if fewer than the specified maximum number of vehicles are simultaneously slipping slots there. If no vacant slots are available, the vehicle desiring to demerge is instead advanced forward on the main line. The vehicle will then be rerouted to its destination via the best route from its current position downstream of the demerge switch. A vehicle denied access to an interchange ramp is said to have been aborted. Each abort is recorded by the computer. In the case of T-intersections, such as ramp 15 in the network of Figure 1, the line from station 8 to station 10 is run synchronously. Slot slipping is permitted downstream of the merge from station 9. If all slots on the line downstream of station 9 are filled, only vehicles going to station 9 are allowed to enter the line upstream of station 9. All other vehicles are rerouted at the interchange upstream from station 9.

Nominal vehicle routing and the rerouting of aborted vehicles is managed on the basis of a minimum-path table. At each intersection, the destination of the vehicle is interrogated, and the appropriate entry in the minumum-path table is examined. The minumum-path table is an NxN array, where N is the number of stations in the network. The element located in row i and column j of the indexed array is the number of the next station which should be passed on the best path from station i to station j.

This "best path" is derived from the expected average origin-destination demand matrix using a minimum-path algorithm with a link capacity constraint. This table is updated periodically. Each interchange is associated with the station just upstream from it, so that the best path is determined directly from the vehicles present location and its destination.

At each update, demands for service are also generated. These demands, while random in nature, have a mean which corresponds to the specified input data. The actial demand arrivals are generated by a Monte Carlo process. Because of the nature of PRT, passenger demands will probably not occur as a Poisson distribution. Passengers will arrive at stations in small groups with the intention of riding together. Each group, however, represents a vehicle demand, and it is reasonable to assume that the arrival of each group representing a vehicle demand is a Poisson process. A table of average vehicle demands per hour from each origin station to every destination station is specified apriori as an input to the simulation program. The actual demand arrival pattern is a Poisson distribution with a mean equal to the given average demand.

Separate subprograms also provide data for the simulation program. One of these subprograms is an optimal empty-vehicle shuttling algorithm which minimizes the average total empty-vehicle-trip mileage per hour, while ensuring that the average waiting time at each station in the network is less than a specified value. This subprogram uses the given demand data to compute an empty-vehicle flow matrix, e_{ij}, which represents the average flow of empty vehicles from station i to station j per hour. A description of this algorithm is given in Appendix A of this paper. An estimate of the required fleet size can be computed with this information. The fleet size is determined from the formula:

$$\text{Fleet Size} = \sum_{j=1}^{N} \sum_{i=1}^{N} t_{ij}(d_{ij} + e_{ij})$$

where t_{ij} is the minimum-path travel time between stations i and j, and d_{ij} and e_{ij} are the full and empty vehicle demands per hour from station i to station j, respectively. Vehicle shuttling requirements, and in turn the fleet size, are influenced by two station parameters, the number of berths in the station, and the upper bound on the average passenger waiting time. Fleet requirements will increase for a fixed vehicle demand matrix if the specified average waiting time at stations is decreased, since this will require an increase in empty-vehicle shuttling activity. In the present computer model, "waiting time" includes only the time spent waiting for a vehicle. Passenger boarding times and intra-station vehicle movement times have not been included. A more detailed model of station operations is being incorporated into the simulation program at the present time.

Vehicle demands are generated at each station by a two-stage Monte Carlo process. The first stage of the process determines whether a vehicle demand has occurred at the station during the update time interval. The second stage of the Monte Carlo process determines the destination of the full or empty vehicle departing from the station. Vehicle dispatching from a station proceeds in the following way at each update time:

(1) If there is a vehicle-demand queue and an empty vehicle in the station, that vehicle is dispatched to a destination chosen at random on the basis of the origin-destination matrix.

(2) If there is no demand queue, and there is an excess empty vehicle in the station, the vehicle is dispatched at random according to the empty-vehicle demand matrix.

The simulation was written in FORTRAN and was designed to run on the CDC 6400 time sharing facility at the University of Minnesota. Running time and storage requirements were significantly reduced by using word-packing on an extensive scale. Each slot on the guideway was represented by six bits of a sixty-bit computer word. Thus, one computer word represents ten slots of guideway. At each time step, the sixty-bit word is shifted six bits to the right, with the right-hand six bits transferred to the left end of the next word. Several indicators of system performance are listed at the end of each run. These include:

(1) The number of passengers waiting at each station at the end of a given time period.

(2) The average passenger waiting time at each station over a given period.

(3) The number of aborts encountered at each interchange and station in the network.

(4) The total vehicle flows on all the links in the network.

(5) The average time delay per vehicle at each interchange in the system.

These are some of the more important measures of system performance. Information on maximum waits and delays, as well as the distribution of waiting times may also be obtained.

A Numerical Example

A network of 23 stations and 17 interchanges was set up to examine some of the command and control problems associated with a high-capacity small-vehicle automated transit system. A scale drawing of the network is shown in Figure 1. The empty-vehicle shuttling algorithm described in Appendix A was also inplemented here so that its effectiveness could be determined in an operating situation with random passenger demands.

In order to test a high-demand situation, a vehicle-demand level which would require about fifty-percent of the total slots on the main lines to be occupied was considered. At one-second headway, this represented a fleet size of about 1100

vehicles operating on the network of Figure 1. The origin-destination demand matrix was constructed by first postulating a set of productions and attractions for each station in the network. Care was taken to chose values which might reasonably correspond to an actual non-symmetric denand distribution. The values used for these productions and attractions are presented in the bar-graph at the bottom of Figure 2. The demand distribution is clearly non-uniform in nature; some stations, such as stations 2, 5, and 10, will require a large influx of empty vehicles to satisfy passenger demand, while others, such as stations 6, 15, and 23, will have a large outflow of empty vehicles. The complete origin-destination matrix was obtained from the data on productions and attractions by using a gravity model with the gravity constant set to zero. Average flow calculations based on this data predicted a fleet requirement of about 1070 vehicles for the network of Figure 1, assuming a one-minute average wait and three-berth stations. (Station configurations in the system actually ranged from two to five berths depending on demand level.) Link flows for this demend level, based on average flow calculations along minimum paths, are given in Table 1. Several of the link flows in this table are in the range of eighty percent of the theoretical line capacity of 3600 vehicles per hour, assuming one-second headways. The average empty-vehicle demand matrix was computed from the origin-destination demand matrix using an upper bound of one minute for the average waiting time at each station. The optimization procedures used to compute this matrix are described in Appendix A. This matrix is required as data for the simulation program; other inputs to the program include the minimum-path table and various network-description arrays.

The network of Figure 1 was first initialized by loading every other slot on the main line with a vehicle whose destination was chosen at random. The simulation was then run for 10,000 time steps and the system state at the end of the run was saved for later retreival. The purpose of this run was to remove the transient effects due to the particular initial conditions used. After this initial run, several 6000-second runs were made, each run using the final state of the previous run as the initial state. The same origin-destination demand matrix was used as the input for each run.

The results of three of these runs are shown in Figure 3. The average waiting times recorded here were taken at real-time intervals 6000 seconds apart. In each case, the average was taken over a 6000-second interval. No parameters were changed in the intervals; all variations in average waiting times between intervals are due solely to the random nature of the passenger demands. The bar-graph of Figure 3 shows a considerable fluctuation of waiting time around the one-minute average wait specified at each station. By comparing the three data samples, it is apparent that long waiting times are not associated

with any one particular station. A station with a high average wait in one time interval may have a low average wait over the next interval. It should be noted here that the very low waits consistently found at some of the stations in the system are not due to random effects. For example, stations 6, 7, and 23 each have average waits below 15 seconds for each time interval. At these stations, the arrival of full vehicles is significantly higher than the vehicle demand rate at the station. The vehicle-shuttling procedure is designed to assure that no station will have a theoretical average wait, under steady-state conditions, which will exceed one minute. Lower average waits are possible at stations where passenger arrivals exceed service demands.

The large fluctuations in average wait at some of the stations in the system, as demonstrated in Figure 3, can be partially explained by an analysis of an individual station. The theoretical waiting-time distribution curve for a typical station (three berths, 300 vehicle demands per hour, one minute average) is shown in the graph at the bottom of Figure 3. This curve represents the probability that a wait will exceed a given time interval. The variance of such a distribution is large; in fact, except for a one-berth station, the variance will exceed the mean. This would indicate that the average of a discrete sample size could deviate significantly from the mean of the distribution. To test this hypothesis, a single-station simulation was performed. Results show that for a vehicle supply which would theoretically produce an average wait of one minute, five percent of the sample averages of waiting time were more than twice as large as the theoretical mean, the averages being taken over a 6000-second interval.

While open-loop vehicle shuttling policies can do little to change this variance in waiting time, a "maximum" waiting time can be guaranteed with high probability. For example, if vehicle shuttling is determined so that the probability of a person waiting more than T seconds is less than .001, the passenger's actual waiting time may vary considerably, but his maximum wait is closely controlled. The use of such a policy is described in Appendix A. For the range of passenger demands considered in this example, with an upper bound on the waiting time of one minute, calculations show that the probability that the wait will exceed five minutes is less than .001. Closed-loop vehicle-shuttling policies, in which the destination of an empty vehicle is determined by the present state of the system, may be advantageous in reducing the variance in waiting times at a station.

The distribution of waiting times for a typical station in the system is shown in Figure 4. In this figure, passenger waiting time is divided into ten-second intervals. Each bar represents the percentage of passengers whose waiting times are within ten seconds of the value represented on the abscissa. The two bar-graphs represent data for a single station collected over two consecutive 6000-second intervals. In the first

interval, the average wait at the station was found to be about 40 seconds. In the second interval the average wait was 85 seconds. In neither case did the maximum wait for any passenger exceed 3.5 minutes; the longer waiting time is caused by a correspondingly larger number of passenger waits in the two-to-three minute range. Data on passenger queues at this station shows that a small but significant passenger queue developed early in the time interval which corresponded to the larger average wait.

The table at the top of Figure 2 demonstrates that the prescribed open-loop vehicle shuttling policy is being carried out by the system. A close agreement between predicted and actual empty-vehicle flows is achieved at each station. In fact, a slight over-supply of vehicles is indicated here, since in most instances the actual empty-vehicle flow exceeds the predicted value based on steady-state calculations. This close agreement tends to support the conclusion that the variance in waiting times shown in Figure 3 is a local effect due to random fluctuations in vehicle supply and demand at individual stations, and not some peculiarity of the operating strategy.

Figure 5 shows curves of cumulative vehicle and passenger arrivals at various stations. The stations were selected to give a cross-section of demand patterns, but all represent demand situations where some vehicle shuttling is necessary. At station 4, for example, there are about 400 passenger departures and 200 arrivals per hour. The deficit is made up by empty-vehicle shuttling from station 7. In all cases, the total vehicle supply to each station exceeds the vehicle demand. This must always occur for steady-state operation.

Vehicle rerouting due to interchange conflicts was kept to very low values but not entirely eliminated. Some conflicts occurred at interchanges 7 and 8, which can be explained by examining the values of average link flows in Table 1. The total flow downstream of interchange 8 was about 75 percent of capacity, and the flow downstream of interchange 7 was about 65 percent of capacity. Interchanges 7 and 8 averaged one and three aborts per hour, respectively. The average delays experienced at interchanges were quite small in all cases. The largest delay, occurring at interchange 4, was an average of 3.5 seconds. In order to insure that no unresolvable conflict would occur at the T-intersections 15 and 4, up to 20 vehicles were permitted to slip slots simultaneously on the links downstream of stations 9 and 4. A provision was also available for rerouting vehicles away from these links if the queues became too long. In this case, rerouting was performed at successive interchanges upstream of the T-intersections. This provision was never required at any time during the simulation, even though flow through the T-intersection 4 was running at 85 percent of capacity.

Conclusion

A simple management strategy using a quasi-synchronous control scheme has been shown by computer simulation to operate well under a high-demand situation. By managing average line flows in the system, conflicts at interchanges can be reduced to very low values. The empty-vehicle shuttling algorithm employed here is shown to redistribute vehicles efficiently in the network and provide for low average waiting times at stations. Numerical results indicate, however, that the variance in waiting time can be significant due to the nature of the stochastic processes involved. Closed-loop vehicle shuttling policies may be required to significantly reduce this variance.

References

(1) Hadju L. P., Gardiner K. W., Tomura H. and Pressman G. L. "Design and Control Considerations for Automated Ground Transportation Systems". Proc. IEEE 56, 493-513. (1968)

(2) Godfrey M., "Merging in Automated Transportation Systems", Ph.D. Thesis, MIT, 1968.

(3) Boyd R. K. and Lukas M. P., "How to Run an Automated Transportation System.", IEEE Trans. Sys., Man, and Cybernetics, SMC-2, 331-341. (1972)

(4) TRW Systems Group, "A Study of Synchronous Longitudinal Guidance as Applied to Intercity Automated Highway Networks", TRW Report 06818-W66-R0-00. Sept, 1969

(5) Roesler W. J., Waddell M. C., Ford B. M., and Davis E. A., Operating Strategies for Demand-Activated ACGV Systems, Volume 1 and Volume 2, Applied Physics Laboratory, August 1971 and March 1972.

(6) Munson V. A. "Quasi-Synchronous Control of High-Capacity PRT Networks", in Personal Rapid Transit., Ed. J. E. Anderson, J. L. Dais, and A. L. Kornhauser, University of Minnesota (1972)

(7) G. Hadley, Linear Programming, Addison-Wesley Publishing Co., Reading, Mass. (1962) Pages 351-359.

Acknowledgment

This work was supported by a Research and Training Grant from the Urban Mass Transit Administration (UMTA).

Table 1 Link Flows Derived From Average Flow Calculations

Orig Stat	Term Stat	Flow	Orig Stat	Term Stat	Flow	Orig Stat	Term Stat	Flow
1	2	3025	16	17	2299	9	10	1126
2	8	1501	17	19	986	11	12	1230
3	6	1304	18	15	1112	12	9	214
5	1	1316	19	20	990	14	16	897
6	22	1114	20	5	1316	15	7	1327
7	4	1485	21	23	785	17	18	1312
8	10	1502	23	14	698	18	14	199
10	11	2629	2	3	1524	20	21	1161
11	13	1400	3	4	221	21	22	375
12	7	1016	4	1	1706	22	20	1489
13	16	1401	6	23	1045	23	15	1130
15	9	912	7	6	857			

Appendix A
An Empty-Vehicle Shuttling Algorithm

Consider a station with X vehicle storage berths, a demand rate of p vehicle demands per second, and a vehicle arrival rate of q vehicles per second. If the average waiting time for a passenger is to be less than T seconds, then an analysis based on queuing theory requires that

$$qT \geq \lambda^X/(1-\lambda) \qquad (\lambda = p(1-q)/q(1-p)) \quad (1)$$

This result is obtained by modeling the vehicle demand and vehicle arrival probabilities as Poisson distributions. Analysis is simplified by assuming that vehicle dispatching from a station is instantaneous, and thus a passenger queue and an empty-vehicle queue cannot exist simultaneously. It is also assumed here that the probability of two passenger demands occurring in a one-second interval is very small. The same analysis shows that a passengers wait will exceed T* seconds with probability ϵ or less if

$$\epsilon \geq \lambda^X \exp(-q(1-\lambda)T^*) \qquad (2)$$

Depending on the station operating policy which is chosen, relations (1) or (2) will provide a lower bound on the vehicle flow requirements into the station to satisfy that policy. The constraint on vehicle inflow may be written as

$$q \geq F(p,X,T) \qquad (3)$$

where T is the average waiting time if (1) is used, or the "maximum" wait if (2) is used. If the full vehicle flow into the station is not sufficient to satisfy (3), then empty vehicles must be dispatched to the station to meet this constraint. If e_{ij} represents the flow of empty vehicles from station i to station j per hour, then

$$\sum_{j=1}^{N} e_{ji} = (3600)*\begin{cases} F(p_i,X,T) - a_i & \text{for } F(p_i,X,T) \geq a_i \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{j=1}^{N} e_{ij} = (3600)*\begin{cases} p_i - F(p_i,X,T) & \text{for } F(p_i,X,T) \geq a_i \\ p_i - a_i & \text{otherwise} \end{cases} \qquad (4)$$

where $F(p_i,X,T)$ is defined by relations (1) or (2) and p_i and a_i are the arrival rates per second of vehicle demands and full vehicles respectively. The relations of (4) consist of 2N equations and N^2 unknowns, and may be therefore considered as the constraints of an optimization problem for N $\geq$ 2. The objective function chosen for minimization in this paper was the total empty-vehicle mileage traveled per hour. The problem can be posed in the following way:

Under the constraints of (4) and $e_{ij} \geq 0$,

$$\text{Minimize } Z = \sum_{j=1}^{N} \sum_{j=1}^{N} c_{ij}\, e_{ij}$$

where c_{ij} is the distance between station i and station j. A number of efficient methods are available for the solution of this problem (7).

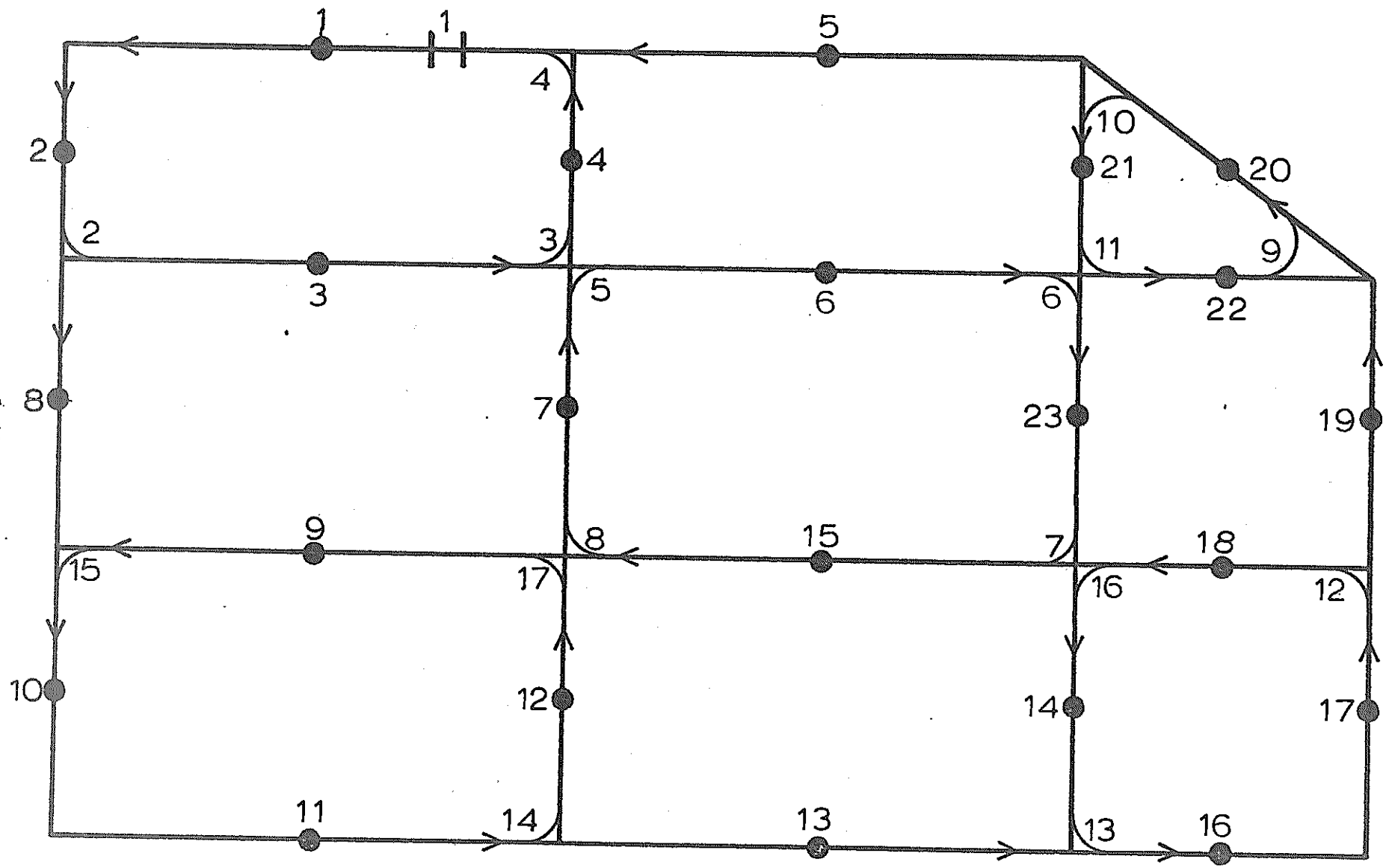
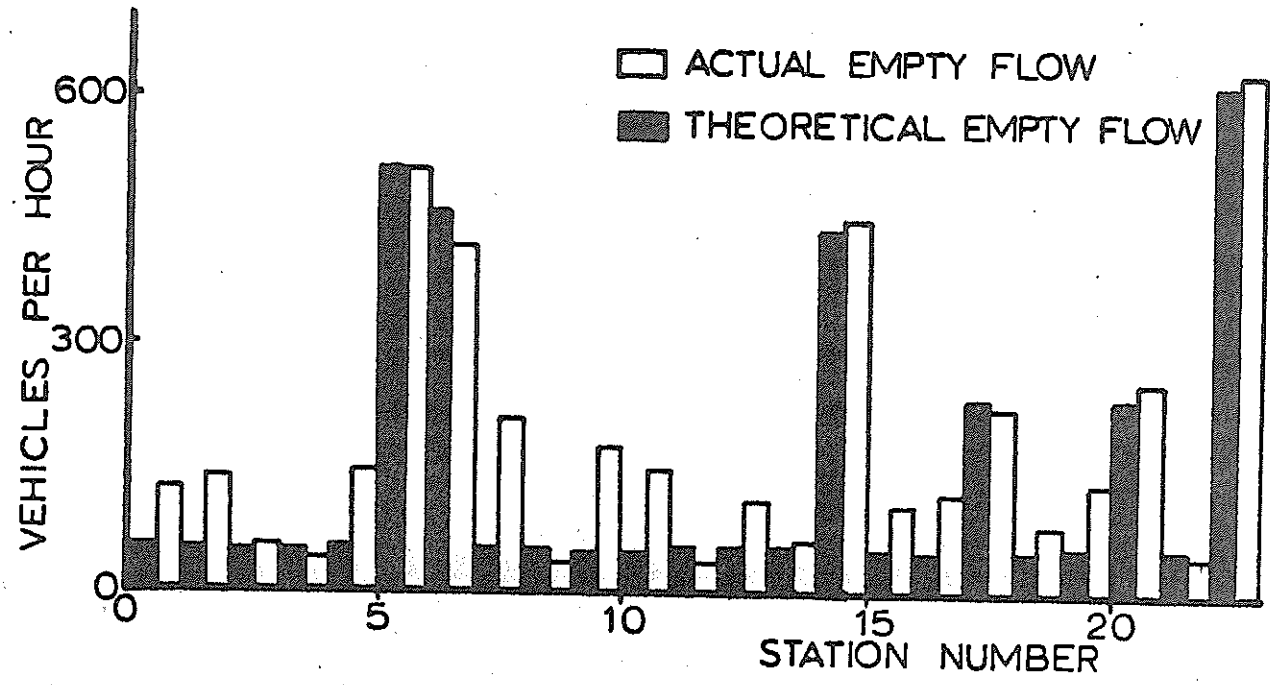


FIGURE 1 TEST NETWORK


ACTUAL EMPTY FLOW
THEORETICAL EMPTY FLOW
VEHICLES PER HOUR
600
300
0
0
5
10
15
20
STATION NUMBER

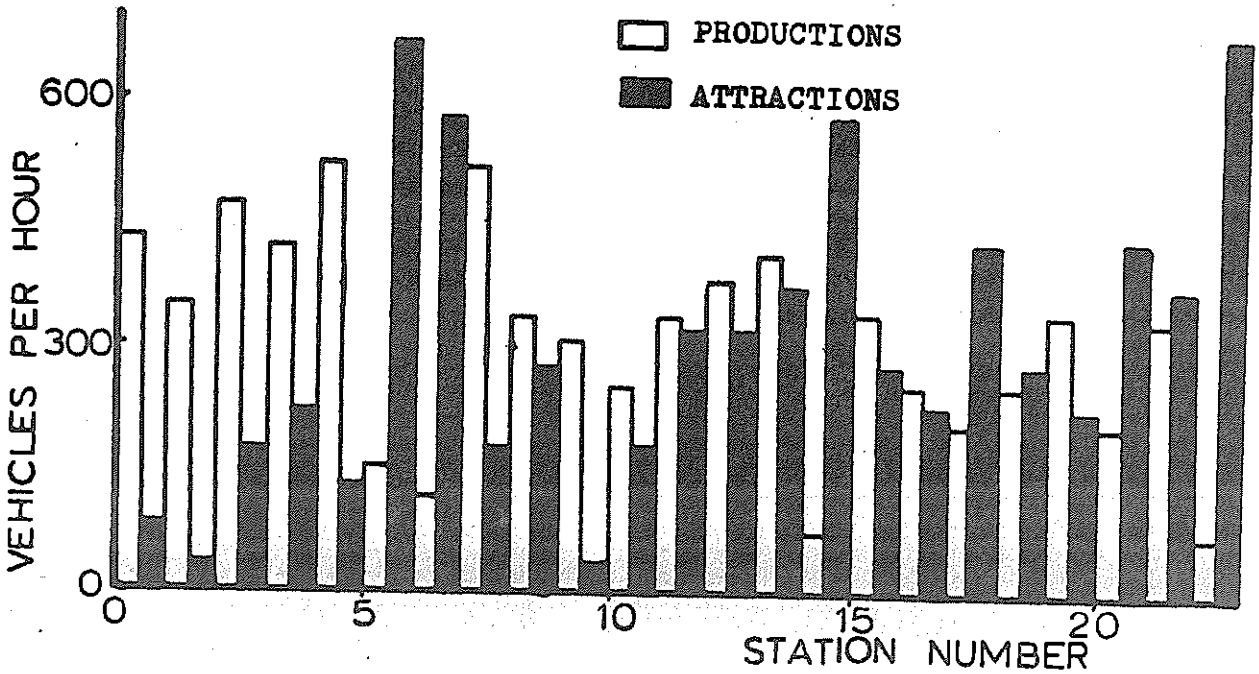
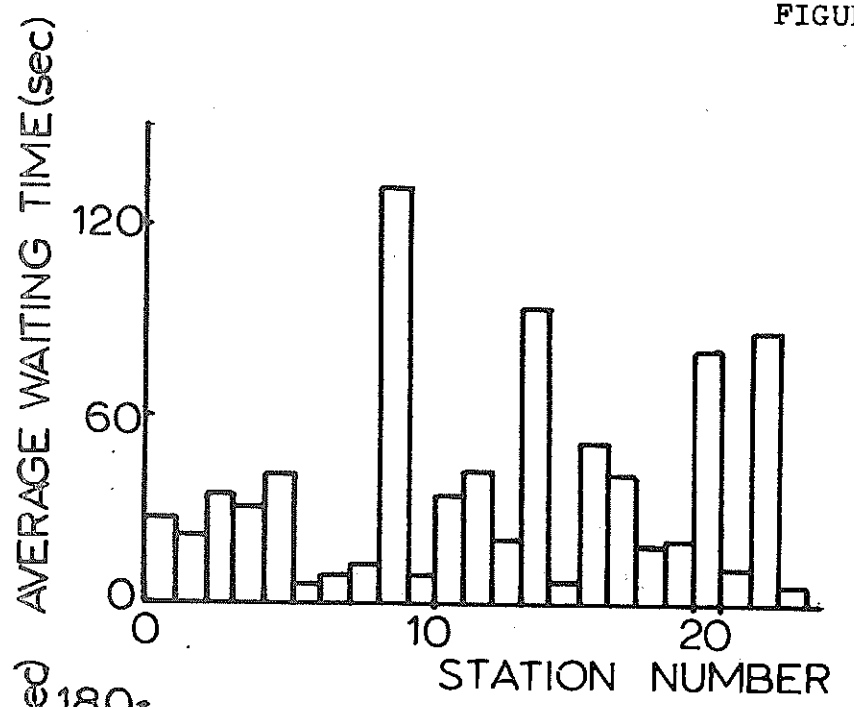
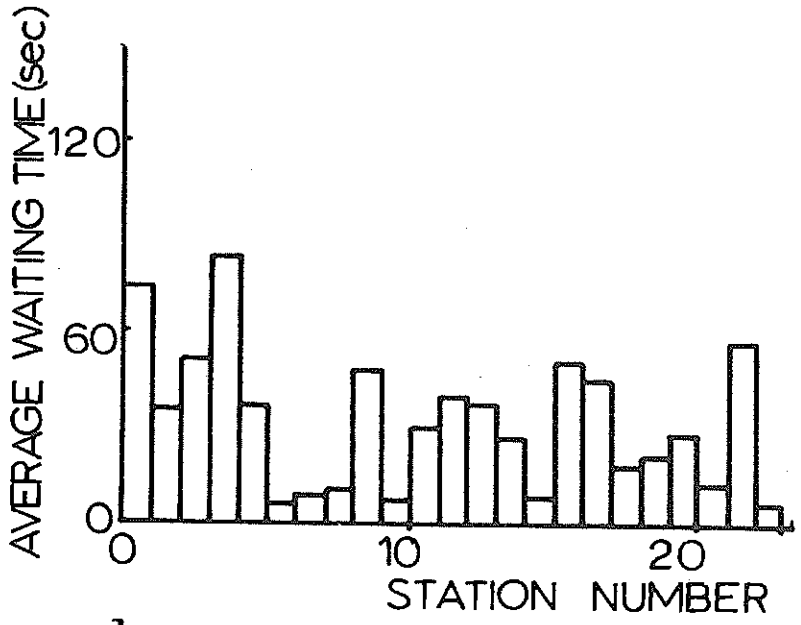
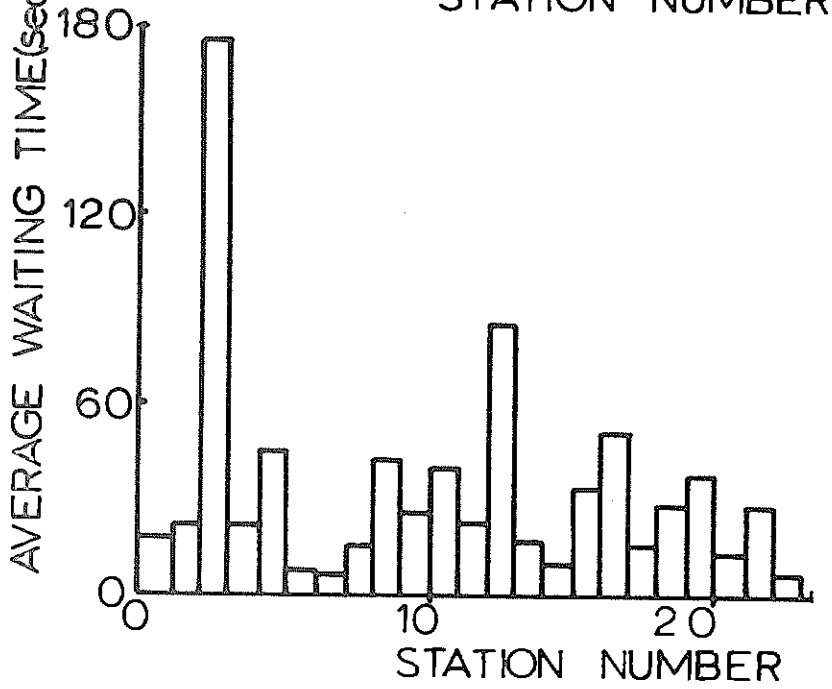

PRODUCTIONS
ATTRACTIONS
VEHICLES PER HOUR
600
300
0
0
5
10
15
20
STATION NUMBER

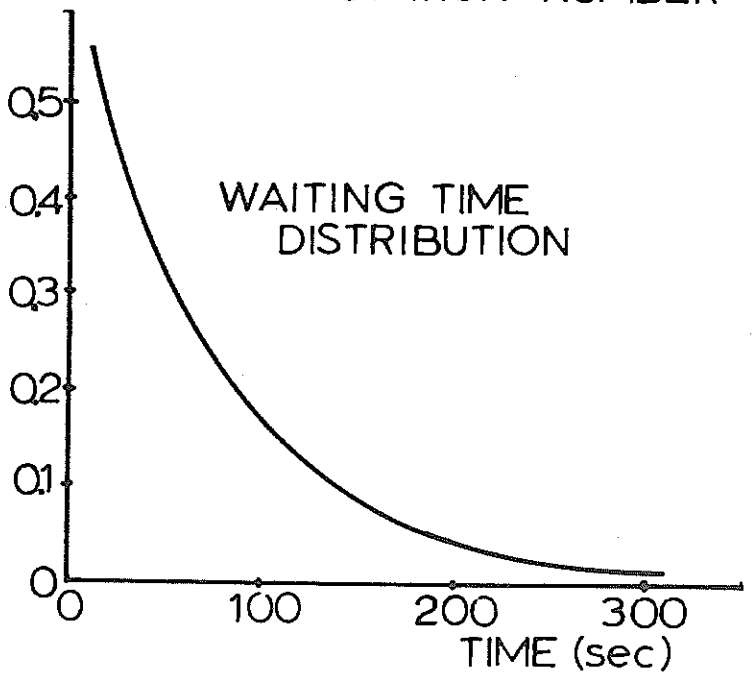
FIGURE 2

FIGURE 3


AVERAGE WAITING TIME(sec)
120
60
0
0
10
20
STATION NUMBER


AVERAGE WAITING TIME(sec)
120
60
0
0
10
20
STATION NUMBER


AVERAGE WAITING TIME(sec)
180
120
60
0
0
10
20
STATION NUMBER


WAITING TIME
DISTRIBUTION
0.5
0.4
0.3
0.2
0.1
0
0
100
200
300
TIME (sec)

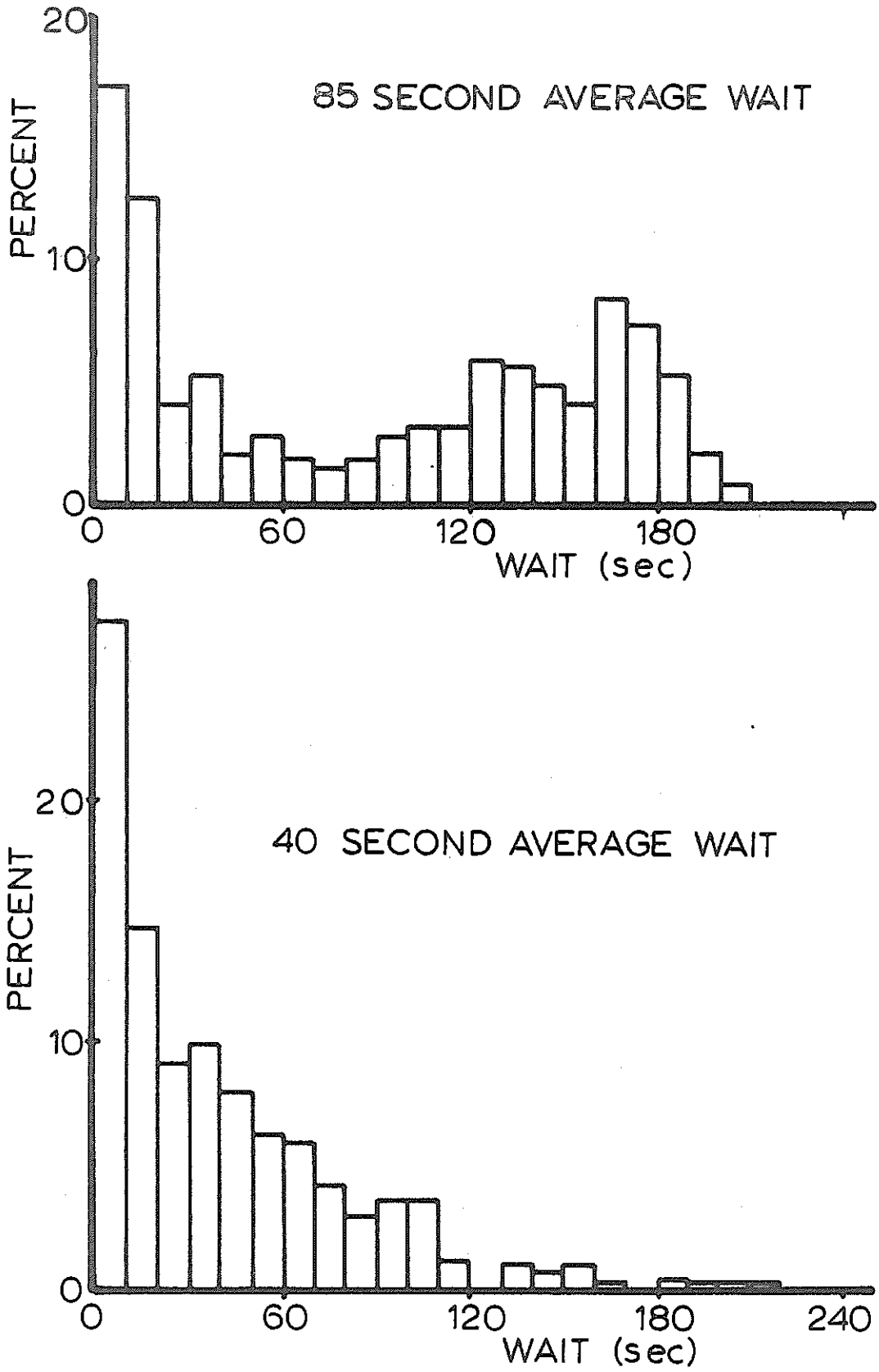
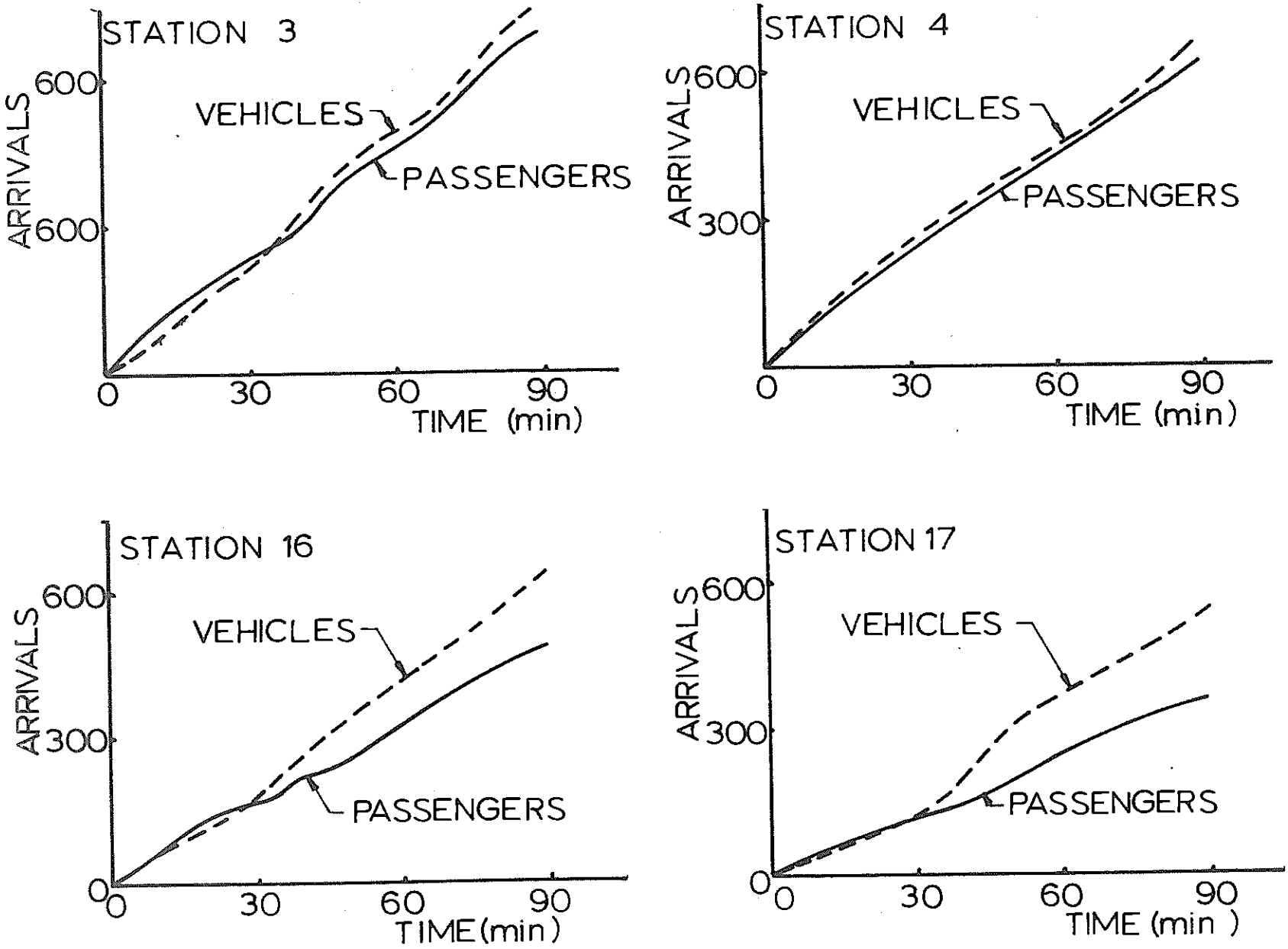

85 SECOND AVERAGE WAIT
PERCENT
20
10
0
0
60
120
180
WAIT (sec)
40 SECOND AVERAGE WAIT
PERCENT
20
10
0
0
60
120
180
240
WAIT (sec)

FIGURE 4

FIGURE 5


STATION 3
ARRIVALS
600
600
VEHICLES
PASSENGERS
0
30
60
90
TIME (min)
STATION 4
ARRIVALS
600
300
VEHICLES
PASSENGERS
0
30
60
90
TIME (min)
STATION 16
ARRIVALS
600
300
0
VEHICLES
PASSENGERS
0
30
60
90
TIME(min)
STATION 17
ARRIVALS
600
300
0
VEHICLES
PASSENGERS
0
30
60
90
TIME (min)

Design of Optimal Feedback Systems for Longitudinal Control of Automated Transit Vehicles*

W. L. Garrard and A. L. Kornhauser

ABSTRACT

Optimization theory is applied to the design of feedback control systems for high-capacity automated transit systems. The resulting controllers are shown to keep headway and velocity errors small without causing passenger discomfort. Excellent dynamic response is achieved during normal mainline operation, merging and demerging, manuevering and emergency stopping. Useful design charts are presented and the effects of the dynamics of the propulsion system are considered. It is concluded that optimization theory is a useful tool in the design of longitudinal control systems for automated transit systems.

Design of Optimal Feedback Systems for Longitudinal Control of Automated Transit Vehicles

W. L. Garrard and A. L. Kornhauser**

1. INTRODUCTION

Almost all new urban transportation systems involve the use of automatically controlled vehicles. Although physical characteristics vary from system to system, the control problems encountered are similar. In many systems, such as personal rapid transit (PRT), dual-mode, and automated freeways, the number of passengers per vehicle is small, thereby short headway are necessary for high capacity. This requires a versatile and efficient control system which must maintain the proper spacing between vehicles without causing passenger discomfort. The control system must be reasonably economical to implement, adaptable to merging and demerging from off-line stations and maneuvering at interchanges, simple enough to insure reliability, and suitable for use in emergency situations (Hajdu, et al, 1968).

Two control philosophies for automated transit systems have evolved. The first is the moving-reference or vehicle-following concept (Athans et al, 1966, Fenton et al, 1971, Brown, 1971, Garrard, et al, 1972, Hesse, 1972). In systems based on this concept, each vehicle receives information directly from other vehicles on the

---

* Department of Aerospace Engineering and Mechanics and Center for Control Science, University of Minnesota, Minneapolis, Minnesota.

** Assistant Professor, Associate Director, Transportation Program, Department of Civil and Geological Engineering, Princeton University, Princeton, New Jersey.

guideway or merging ramps. Control decisions are based on this information. The second control philosophy is the fixed-reference or moving-slot concept (Boyd and Lukas, 1972, Munson, 1972, Wilkie, 1970, Whitney and Tomizuka, 1972). In systems based on this concept, vehicles do not communicate directly with one another, but instead follow a hypothetical slot moving along the guideway at the nominal line velocity. Fixed-reference systems differ from the moving-reference systems in that the position of the vehicle is determined with respect to the guideway rather than with respect to the other vehicles within the system.

A number of investigators have concluded that fixed-reference systems are superior to moving-reference systems in high-capacity transit systems in which vehicles must merge and demerge from off-line stations and maneuver at interchanges. (Boyd and Lukas, 1972, Munson, 1972, Wilkie, 1970, Whitney and Tomizuka, 1972). Fixed-reference systems appear to be superior for the following reasons:

1. Communication problems are reduced since intervehicular communication is not necessary.
2. Merging strategies are easily implemented.
3. Possible "shock-wave" type instabilities due to maneuvering are avoided.

In the present study optimization theory is applied to the design of fixed-reference feedback control systems for high capacity transit systems such as PRT and dual-mode. The resulting optimal controllers are shown to keep headway and velocity errors small without causing passenger discomfort. Excellent dynamic response is achieved during normal mainline operation, merging and demerging, maneuvering, and emergency stopping. The effects of the dynamics of the propulsion system on the response of the vehicle are considered. Useful design charts are presented and applied to the design of a longitudinal control system for a high-capacity PRT system.

2. VEHICLE DYNAMICS AND THE CONTROL SYSTEM MODEL

2.1 Equation of Motion

The differential equation describing the longitudinal motion of the transit vehicle is

$$M\frac{dV}{dt} = -F_D(V,V_W) + F - Mg \sin\theta - F_M \qquad (1)$$

where:

M = mass of the vehicle

V = velocity of the vehicle

V_W = velocity of the wind (positive for a head wind)

F = propulsive force

g = gravitational acceleration

θ = slope of the guideway

F_D = aerodynamic drag

F_M = mechanical resistance

The aerodynamic drag is

$$F_D = C_D(V + V_W)^2 \tag{2}$$

where C_D is a drag "coefficient." Furthermore, the propulsive force is assumed to be governed by

$$\frac{dF}{dt} = -(\frac{1}{\tau})F + Gi \tag{3}$$

where:

τ = time constant of the propulsion system

i = control input to the propulsion system

G = gain constant of the propulsion system.

That is, the propulsion system is modeled as a first-order lag. The error, e, is defined as

$$e = X - X_c \tag{4}$$

where

X = the actual position of the vehicle

X_c = the desired or command position of the vehicle.

Since $V = \frac{dX}{dt}$, (1) can be re-written in terms of the error as

$$\frac{d^2e}{dt^2} = -\frac{C_D}{M}(\frac{de}{dt} + \frac{dXc}{dt} + V_W)^2 + \frac{F}{M} - g \sin\theta - \frac{d^2X_c}{dt^2} - \frac{F_M}{M} \cdot \tag{5}$$

2.2 Nondimensional Formulation

For purposes of generality, the system equations will be non-dimensionalized. The following non-dimensional variables will be used:

$v = \frac{T^2 Gi}{MV_N}$, non-dimensional control input to the propulsion system

$w = \frac{Vw}{V_N}$, non-dimensional headwind velocity

$y = \frac{e}{H}$, non-dimensional error

$y_c = \frac{X_c}{H}$, non-dimensional command position

$\sigma = \frac{t}{T}$, non-dimensional time

$f = \frac{TF}{MV_N}$, non-dimensional propulsive force

$\cdot = \frac{d}{d\sigma}$, derivative with respect to non-dimensional time

where:

V_N = nominal velocity of the vehicles on the main guideway

H = nominal nose to nose distance between vehicles on the main guideway

T = nominal time headway between vehicles, $T = \frac{H}{V_N}$.

The resulting system equations are

$$\ddot{y} = -\left(\frac{C_D}{M}\right) H(\dot{y} + \dot{y}_c + w)^2 + f - \frac{Tg \sin \theta}{V_N} - \ddot{y}_c - \frac{F_M T}{MV_N} \qquad (6)$$

$$\dot{f} = -\frac{T}{\tau} f + v \qquad (7)$$

During normal mainline operation the vehicle will operate at near nominal velocity, and linearization of (6) and (7) is legitimate. Furthermore $\ddot{y}_c$, the commanded acceleration, will be zero, and $\dot{y}_c$, the commanded velocity, will be unity. The

resulting linearized equations of motion are

$$\ddot{y} = -\left(\frac{2C_D}{M}\right) H (1 + w)\dot{y} + f - d$$

$$\dot{f} = -\frac{T}{\tau} f + v \qquad (8)$$

$d = \frac{Tg \sin \theta}{V_N} + \frac{F_M T}{V_N M} + \left(\frac{C_D H}{M}\right)(w + 1)^2$, the non-dimensional disturbance force.

The velocity of the headwind is not known a priori; however, its average value will in general be zero. Thus the coefficient of $\dot{y}$ in (8) will be approximated by its average value $\frac{2C_D H}{M}$. As will be shown subsequently, a feedback controller designed on the basis of the above assumptions provides excellent dynamic response for mainline operation with headwinds three times the nominal vehicle velocity; for merging and demerging from off-line stations; for maneuvering at interchanges; and for emergency control.

3. SYNTHESIS OF THE OPTIMAL FEEDBACK CONTROL SYSTEM

3.1 State Variables

The vehicle and propulsion system have been modeled by a set of linear differential equations with constant coefficients. If the state variables are selected properly, it is possible to use optimization theory to design a feedback control system which will keep headway and velocity errors small without causing passenger discomfort (Athans, 1971). The appropriate state variables for this problem are headway error, velocity error,

acceleration error, and rate of change of propulsive force. Headway and velocity error are obvious choices for state variables; however, acceleration error and rate of change of propulsive force are less obvious candidates. The reasons for selecting these two quantities as state variables will be discussed in detail below.

The state variables are $x_1 = y$, the non-dimensional headway error; $x_2 = \dot{y}$, the non-dimensional velocity error; $x_3 = \ddot{y}$, the non-dimensional acceleration error; and $x_4 = \dot{f}$, the rate of change of the non-dimensional propulsive force. The control variable is $u = \dot{v}$, the rate of change of the non-dimensional input to the propulsion system. For purposes of control system design, the non-dimensional disturbance force, d, is assumed constant (the headwind is constant or the vehicle is ascending or descending a constant slope). Using these definitions and assumptions, the equations of motion of the vehicle can be written in vector-matrix form as

$$\dot{\underline{x}} = A\underline{x} + \underline{b}u \tag{9}$$

where

$$\underline{x}^T = [x_1 \; x_2 \; x_3 \; x_4],$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{2C_D H}{M} & 1 \\ 0 & 0 & 0 & -\frac{T}{\tau} \end{bmatrix},$$

and

$$\underline{b}^T = [0 \; 0 \; 0 \; 1].$$

The control u will be selected in such a manner as to drive the state variables to zero. The state variables x_1 and x_2 represent the position and velocity errors and the necessity of driving these quantities to zero is clear. The state variables x_3 and x_4 represent the acceleration error and the rate of change of propulsive force. The acceleration error must be zero if the position and velocity errors are to remain zero; furthermore, in order to achieve zero acceleration error, the propulsive force must equal the disturbance force. Since the disturbance force is assumed to be constant, the propulsive force must also approach a constant value, and the derivative of the propulsive force, x_4, must approach zero.

3.2 The Performance Index

The error equations have been formulated in the standard notation of optimal control theory; however, in order to apply this theory a mathematical criterion for the measurement of system performance is necessary. A quadratic performance index is proposed. This index is

$$J = \frac{1}{2} \int_0^\infty (q_1 x_1^2 + q_2 x_2^2 + q_3 x_3^2 + q_4 x_4^2 + r u^2)\, dt. \qquad (10)$$

The optimal feedback control problem is to determine the control, u, as a function of the state variables in such a manner as to minimize J. It can easily be shown that the control which minimizes J drives the state variables to zero (Lee and Markus, 1967).

The performance index J is the integral of the weighted sum of the position error, velocity error, acceleration error, derivative of the propulsive force, and the control effort. Since the drag coefficient is small, the derivative of the propulsive force is very nearly equal to the jerk. The performance index penalizes large position and velocity errors and the control which minimizes J should result in a system in which these errors are kept small. The performance index also penalizes large acceleration errors and jerks. These two variables affect passenger comfort and the control which minimizes J should also result in a system in which passenger discomfort is minimized. The rate of change of the control input to the propulsion system, u, must be included in the performance index in order to obtain the optimal control in feedback form.

It is of course possible to formulate many other performance indices which include system error and passenger comfort; however, use of a quadratic performance index as given in (10) permits determination of the optimal control in feedback form. This is one of the few classes of optimization problems in which the optimal feedback control can be found (Lee and Markus, 1967). In addition the optimal feedback control is linear with constant gains. Such a controller is easy to implement.

The feedback control which minimizes J is

$$u = -r^{-1} \underline{b}^T K \underline{x} \qquad (11)$$

where K, the optimal gain matrix, is the symmetric, positive-definite solution of the matrix Ricatti equation

$$KA + A^TK - K\underline{b}^Tr^{-1}\underline{b}K + Q = 0 \qquad (12)$$

where $\underline{\underline{Q}} = \begin{bmatrix} q_1 & 0 & 0 & 0 \\ 0 & q_2 & 0 & 0 \\ 0 & 0 & q_3 & 0 \\ 0 & 0 & 0 & q_4 \end{bmatrix}$.

Iterative methods for the solution of (12) allow rapid determination of the optimal gain matrix by use of a high-speed digital computer (Kleinman, 1968).

From (12) the optimal control in terms of the actual error variables is

$$u = -\frac{K_{14}}{H}\dot{e} - \frac{K_{24}}{V_N}\dot{e} - \frac{K_{34}T}{V_N}\ddot{e} - \frac{K_{44}T^2}{MV_N}\dot{F} \qquad (13)$$

where the K_{ij}'s are the elements of the optimal gain matrix K. A block diagram of the vehicle and control system is shown in Figure 1.* It can be seen that the input to the propulsion system is proportional to the headway error, the derivative of the headway error, the integral of the headway error, and the propulsive force. The optimal control (13) is similar to the control derived by Whitney and Tomizuka (1972) using classical techniques and by Wilkie (1970) using optimization theory. However, in neither of these studies was the dynamics of propulsion system considered. It should be noted that in all of these systems,

---

* The symbol "s" denotes the Laplace operator.

steady-state errors due to constant biases in the measurements of velocity errors, acceleration errors, and the rate of change of propulsive force are zero.

3.3 Determination of Weighting Matrix Q.

The weighting factors q_1, q_2, q_3, q_4 and r affect the values of the elements of K, the gain matrix.* The values of the elements of the gain matrix in turn affect the dynamic response of the system. The relationship between the weighting factors and the dynamic response of the vehicle cannot be analytically determined. If the weighting factors on headway and velocity error are chosen to be large relative to the weighting factors on the acceleration errors and jerk (rate of change of propulsive force) a system which zeros errors rapidly but gives an uncomfortable ride will result. On the other hand, if the weighting factors on acceleratic error and jerk are chosen to be large relative to the weighting factors on headway and velocity error, the ride will be comfortable but the system will be rather sluggish in reducing the headway error to zero. Thus the designer must determine how the weighting factors affect the dynamic response in order to obtain the proper trade-off between ride quality and adequate control of headway error.

Figs. 2-11 illustrate various system performance characteristics as functions of the weighting factors. Preliminary

---

* It can be shown that $K_{14} = \sqrt{q_1}$; the other gains must be determined numerically, however.

computations indicated that the velocity error should be weighted about ten times the position error ($q_2 = 10q_1$). This resulted in tighter headway control as well as a more comfortable ride than could be obtained by weighting the position error the same as or greater than the velocity error. Since the performance index can be multiplied by a constant without changing the value of the optimal gain matrix, any one of the weighting factors can be arbitrarily set equal to unity. Thus the weighting factor on the control, r, was chosen to be one. The weighting factor on acceleration error, q_3, was set at ten, and the values of the other state-variable weighting factors were varied with respect to q_3. The coefficient $\frac{2C_DH}{M}$ was set at .025, a value typical of vehicles in many PRT and dual-mode systems (Whitney and Tomizuka, 1972, Wilkie, 19

The headway error, acceleration, jerk, maximum power, and the time to reduce the headway error to ten percent of its initial value were plotted versus the ratio of q_1 to q_3. The maximum acceleration and jerk were determined for an initial headway error, $e_o = 0.1$ H and zero headwind. The maximum headway error and maximum power were determined for $e_o = 0$ and a headwind three times the nominal velocity of the vehicle. The information presented in Figs. 2 - 11 is applicable to a wide variety of systems since the quantities plotted are non-dimensional. An example using specific values of V and H is presented in Section 4. In Figs. 2 - 6 the time constant of the

propulsion system was assumed to be one-tenth of the minimum headway time and weighting factors of 10, 1,000 and 10,000 were placed on the jerk.

It is interesting to note in Fig. 2 that in order to obtain maximum headway errors of less than fifty percent, it is necessary to weight the headway error at least ten times the acceleration error. In Fig. 5, it can be seen that for values of the weighting factor on headway error greater that ten, the time to reduce a headway error to ten percent of its initial value is, for all practical purposes, constant and does not depend on the weighting factor associated with the jerk. As would be expected, Figs. 3 and 4 indicate that acceleration and jerk increase as the ratio of headway error to acceleration error increases. Furthermore, the jerk increases at a faster rate than the acceleration and has larger numerical values than the acceleration. Thus jerk, rather than acceleration is the limiting factor in obtaining tight control. Fig. 6 illustrates the maximum power requirements for various values of the weighting factors. The maximum power required is smaller for large values of the weighting factor on the headway error than for small values and is sensitive to the value of the weighting factor on jerk. It can be seen that acceptable performance is obtained for ratios of q_1 to q_3 between one and ten thousand. Unacceptably large headway errors result if q_1/q_3 is less than one and excessive jerk results if q_1/q_3 is greater than ten thousand. Also peak power requirements are not excessive for q_1/q_3 between one and ten thousand.

The time constant of the propulsion system is an important parameter in the design of the longitudinal control system. For large values of this time constant, the response of the vehicle may be oscillatory or even unstable if propulsion system dynamics are not given proper consideration in the design procedure. The system model used in this study includes propulsion system dynamics and the optimal control compensates for time lags introduced by the propulsion system. This is illustrated in Figs. 7 - 11. The dynamic response characteristics of the optimally controlled vehicle are shown for propulsion system time constants equal to the minimum headway time, one-tenth the minimum headway time, and ten times the minimum headway time.

The response characteristics of the vehicle for $\tau = T$ and $\tau = 10T$ are almost indistinguishable. This can be explained by referring to Table I. The gains for position, velocity, and acceleration errors are almost identical for both values of the propulsion system time constant. The gains for the derivative of propulsive force are different. However, if the natural coefficient of the propulsive force, $\frac{T}{\tau}$, is added to the gain for the derivative of the propulsive force is almost the same for $\tau = T$ and $\tau = 10T$. Thus the dynamic response of the optimally controlled vehicle is for all practical purposes invariant for large values of the time constant of the propulsion system.

It is interesting to note that for the same value of q_1/q_3 tighter headway control was maintained with large propulsion system time constants than with small. However, the accelerations, jerks

and peak power were also larger. The time required to reduce a headway error to ten percent of its initial value did not vary a great deal with the various values of the propulsive system time constant.

Figs. 2 - 11 and Table I should be of considerable use in the design of optimal feedback control systems for automated vehicles as the proper values of the weighting factors for a given set of specifications can be easily selected from the figures. Once the weighting factors have been selected, the proper non-dimensional gains can be found in Table I, and the dimensionalized gains can then be determined from (13). Thus a systematic procedure is presented for the design a longitudinal control system which maintains tight headway control without causing passenger discomfort.

	$\tau = .1T, \quad q_2 = 10q_1, \quad q_4 = 10, \quad q_3 = 10$			
q_1/q_3	K_{14}	K_{24}	K_{34}	K_{44}
10^{-2}	0.316	2.469	7.776	1.205
10^{-1}	1.000	5.823	11.665	1.547
1	3.162	14.866	18.824	2.151
10	10.000	40.664	32.340	3.217
10^2	31.624	117.204	58.687	5.079
10^3	100.000	349.970	111.879	8.270
10^4	316.243	1068.284	222.545	13.563
10^5	1000.000	3304.094	457.215	22.012
10^6	3162.434	10299.756	959.351	35.050

Table I - Non-Dimensional Gains for Various Values of Propulsion System Time Constants

	$\tau = 10T, \quad q_2 = 10q_1, \quad q_4 = 10, \quad q_3 = 10$			
q_1/q_3	K_{14}	K_{24}	K_{34}	K_{44}
10^{-2}	0.316	2.090	5.214	4.420
10^{-1}	1.000	5.030	7.534	4.908
1	3.162	13.458	12.674	5.846
10	10.000	38.474	23.816	7.493
10^2	31.624	114.159	47.669	10.165
10^3	100.000	346.134	98.628	14.300
10^4	316.243	1063.84	207.650	20.529
10^5	1000.000	3299.276	441.348	29.787
10^6	3162.434	10294.722	942.977	43.457

	$\tau = T, \quad q_2 = 10q_1, \quad q_4 = 10, \quad q_3 = 10$			
q_1/q_3	K_{14}	K_{24}	K_{34}	K_{44}
10^{-2}	0.316	2.098	5.265	3.640
10^{-1}	1.000	5.049	7.617	4.122
1	3.162	13.486	12.792	5.049
10	10.000	38.512	23.960	6.670
10^2	31.624	114.203	47.827	9.329
10^3	100.000	346.182	98.794	13.446
10^4	316.243	1063.896	207.825	19.661
10^5	1000.000	3299.327	441.516	28.909
10^6	3162.434	10294.773	943.164	42.572

4. DESIGN OF AN OPTIMAL LONGITUDINAL FEEDBACK CONTROL SYSTEM FOR A HIGH-CAPACITY PRT SYSTEM

4.1 System Specifications

A longitudinal control system was designed for a PRT system with the following specifications:

Nominal Mainline Velocity = 50 ft/sec

Minimum Headway Time = 1 sec

Vehicle weight = 3200 lbs

Vehicle Length = 10 ft

Maximum Acceleration in Mainline Operation = 4 ft/sec^2

Maximum Acceleration for Merging and Manuevering = 8 ft/sec^2

Maximum Emergency Deceleration = 25 ft/sec^2

Maximum Jerk in Mainline Operation = 4 ft/sec^3

Maximum Jerk in Merging and Manuevering = 8 ft/sec^3

Maximum Headway Error = 10 ft

Propulsion System Time Constant = 0.1 sec

The above specifications are typical of many proposed high-capacity PRT systems and result in a system with a mainline capacity of 3600 vehicles/hr.

4.2 Selection of the Optimal Feedback Gains

The minimum headway for the system is 50 feet, thus a maximum headway error of twenty feet is allowed. From Fig. 2, it can be seen that the following values of the weighting factors will result in a control system which satisfies this criterion

q_4	q_3/q_1
10	10 or greater
1000	100 or greater
10000	1000 or greater

Table II - Values of weighting factors which satisfy maximum headway error requirement

However the maximum allowable jerk in mainline operation is 4 ft/sec^3 and from Fig. 4* the following values of the weighting factors will result in a control system which satisfies this criterion.

q_4	q_3/q_4
10	10 or less
1000	100 or less
10000	1000 or less

Table III - Values of weighting factors which satisfy maximum jerk requirements

---

* Non-dimensional jerk is obtained by multiplying dimensional jerk by T^2/V and non-dimensional acceleration is obtained by multiplying dimensional acceleration by T/V.

From Fig. 3 it can be seen that the acceleration criterion of 4 ft/sec^2 is satisfied by the weighting factors given in Table III. Thus the trade-off between passenger comfort and tight headway control gives the following allowable values for the weighting factors.

q_4	q_1/q_3
10	10
1000	100
10,000	1000

Table IV - Values of weighting factors which satisfy maximum headway error and maximum jerk requirements

Thus a choice of three sets of weighting factors is available. From Figs. 5 and 6 it can be seen that the response time and peak power requirements are nearly the same for all three sets of gains. Simulations of merging and mainline operations revealed no important differences in dynamic response characteristics for the three sets of weighting factors; consequently, graphical results are only presented for the case in which the weighting factors are $q_4 = 10$ and $q_1/q_3 = 10$. The non-dimensional feedback gains for this case are given in Table I.

2.3 Simulation Results

The nonlinear vehicle dynamics (6) and the dynamics of the control system (7) were simulated on a digital computer, and the following situations were studied:

1. Mainline operation with a suddenly applied headwind of 150 ft/sec (Fig 12).
2. Mainline operation with an initial headway error of 5 ft (Fig 13).
3. An emergency stop with a constant deceleration of 25 ft/sec^2 (Fig 14).
4. Merging from an off-line station following a trapezoidal acceleration profile (Fig 15). and
5. Manuevering or "slot slipping" following a trapezoidal acceleration profile (Fig 16).

It can be seen that the maximum headway error due to a suddenly applied wind gust of 150 ft/sec is 6 ft and the maximum velocity error is 5.5 ft/sec. The maximum jerk for an initial error of 5 feet is 2.5 ft/sec^3 and the maximum acceleration is 1 ft/sec^2, thus the resulting ride is very comfortable. These results are to be expected since the design was based on mainline operating conditions.

During merging, emergency stopping, and slot slipping the vehicle is not operating at mainline conditions. For example, during merging the vehicle starts with zero velocity and accelerates to line velocity following a commanded acceleration profile as

shown in Fig. 15. The differential equations describing the state of the system during merging, emergency stopping and slot slipping, are nonlinear with time varying coefficients. Our control system was designed on the basis of state equations which were linear with constant coefficients. However, as can be seen from an examination of Figs. 14-16, the resulting controller follows the desired profiles very well even though the conditions are vastly different from those encountered during mainline operation. The results of the merging and slot slipping operations are summarized in Table IV.

At end of manuever Headway Error			At end of manuever Velocity Error		Maximum Jerk
	per cent	ft	per cent	ft/sec	ft/sec^3
Merging	.28	0.14	3.90	1.95	10.13
Slot Slipping	-1.03	-.50	.32	0.16	10.64

Table IV - Final Headway and Velocity Errors for Merging and Manuevering.

Thus a controller designed for mainline operation can also be used for other operations such as emergency stopping, merging, and manuevering. This is important since it shows that a linear controller with fixed gains is sufficient for all control operations and thus the complexity of the control system is minimal.

It should be kept in mind that the emergency deceleration simulation does not consider the effect of slippage between the vehicle and the guideway nor the effects of discrete data sampling and noisy measurements. Such effects can significantly degrade the vehicles stopping ability. The simulation does indicate that the controller does follow even the most severly different acceleration profiles from that of mainline operation (zero acceleration). This implies that

the sensitivity of the computed K_{i_j} values to changes in mainline acceleration conditions is very small. The sensitivity of K_{i_j} to variations in the other assumed nondimensional nominal conditions was not investigated.

5. CONCLUSION

From the results presented above, it appears that optimal control theory can be usefully applied to the design of longitudinal control systems for automated transit systems with a wide variety of characteristics. The resulting control systems keep headway and velocity errors small without causing passenger discomfort, and excellent dynamic response is achieved during mainline operation, merging and demerging, maneuvering and emergency stopping. The controllers are linear with constant gains and should be relatively economical to implement and simple enough to insure reliability.

Since the data presented in Figs. 2-11 and Table I are nondimensional, the results are applicable to a wide variety of systems. The designer should find these data useful in selecting the appropriate feedback gains for various system specifications.

In the present study perfect information sensing has been assumed. In practice of course this is never the case; consequently, the problems associated with noisy sensors, incomplete information and sampled data should be investigated. It is felt that optimal filtering theory (Jazwinski, 1970) can be profitable applied to the problem of noisy and incomplete data. Specific difficulty in implementation is expected resulting from problems in measuring $\ddot{e}$ (acceleration error) and $\dot{f}$ (rate of change of propulsive force). If such measurements are available, they will be quite noisy and a

degradation and possible instability in the controller performance may result. Efforts are being made to design an observer for these state variables in an attempt to eliminate the requirement of measuring $\ddot{e}$ and $\dot{f}$. Also the sensitivity of the response of the optimally controlled vehicle to parameter variations such as changes in vehicle weight due to passengers loading or unloading needs further study.

ACKNOWLEDGEMENT

This research was partially supported by grant Minn-URT-3(71) of the Urban Mass Transit Administration of the United States Department of Transportation.

Athans M. Levine W.S., and Levis W.H. (1969). "On the optimal and suboptimal position and velocity control of a string of high speed vehicles". M.I.T. Electronic Systems Laboratory Report ESL-R-291, Cambridge, Massachusetts.

Athans M. (1971). "On the design of P-I-D Controllers using optimal linear regulator theory." Automatica 7, 643-648.

Boyd R.K. and Lukas M.P. (1972). "How to run and automated transportation system". IEEE Trans. Sys., Man, and Cybernetics, SMC-2, 331-341.

Brown, S.J. (1971). "Characteristics of a Linear Regulator Control Law for Vehicles in an Automatic Transit System." Presented at 1971 AIAA Guidance and Control Conference, Stony Brook, N.Y.

Fenton, R.E., Olson K.W., and Bender J.G. (1971). "Advances toward the automatic highway". Proc. HRB 50th Annual Meeting, Washington, D. C.

Garrad W.L., Hand G.R., Raemer R. (1972). "Suboptimal feedback control of a string of vehicles moving in a single guideway." Transpn Res. 6.

Hadju L.P., Gardiner K.W., Tomura H. and Pressman G.L. (1968). "Design and control considerations for automated ground transportation systems." Proc. IEEE 56, 493-513.

Hesse, R. (1972). "German experiences in the planning and development an an automatic cabin taxi system." Proc. Intersociety Conference on Transpn., Washington, D. C.

Jazwinski A. H. (1970). Stochastic Processes and Filtering Theory, Academic Press, New York,

Kleinman, D. L. (1968). "On an iterative technique for Ricatti equation computations." IEEE Trans. Aut. Control AC-13, 114-115.

Lee, E.B. and Markus L. (1967). Foundations of Optimal Control Theory. John Wiley & Son, New York

Munson A.V. (1972). "Quasi-Synchronous control of high-capacity PRT networks". in Personal Rapid Transit, ed. J.E. Anderson, J.L. Dais, W.L. Garrard, and A.L. Kornhauser, Univ. of Minnesota, 325-350.

Whitney D.E. and Tomizuka M. (1972). Normal and emergency control of a string of vehicles by fixed reference sampled-data control. in Personal Rapid Transit, ed. J.E. Anderson, J.L. Dais, W.L. Garrard, and A. L. Kornhauser, Univ. of Minnesota, 383-404.

Wilkie D.F., (1970). "A moving cell control scheme for automated transportation systems." Transpn. Sci., 347-364.

Fig. 1 Block Diagram of Vehicle, Propulsion System, and Feedback Controller

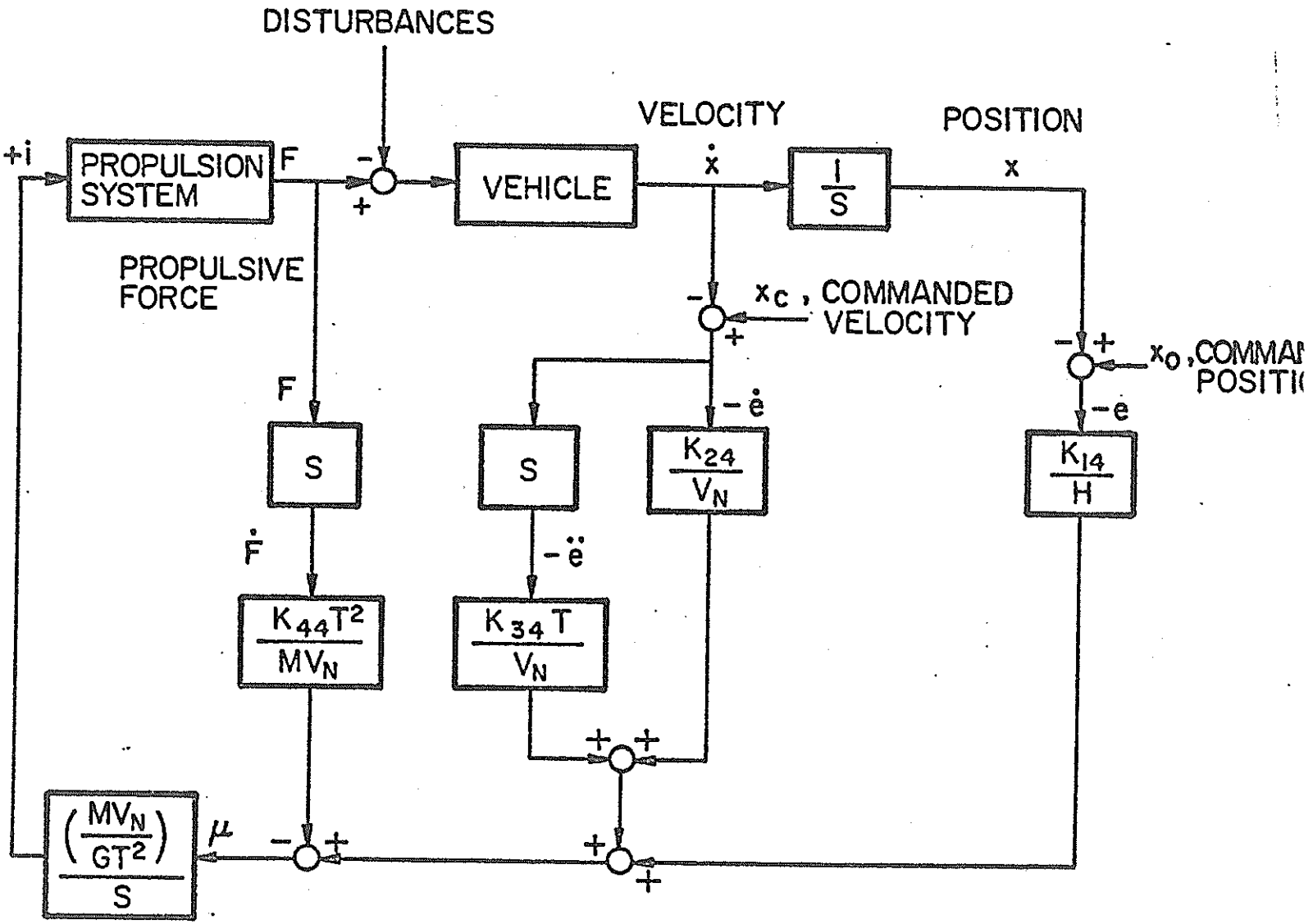

DISTURBANCES
VELOCITY
POSITION
+i
PROPULSION SYSTEM
F
VEHICLE
$\dot{x}$
$\frac{1}{S}$
x
PROPULSIVE FORCE
x_c , COMMANDED VELOCITY
x_o , COMMAI POSITI
F
S
S
$-\dot{e}$
$\frac{K_{24}}{V_N}$
$-e$
$\frac{K_{14}}{H}$
$\dot{F}$
$-\ddot{e}$
$\frac{K_{44}T^2}{MV_N}$
$\frac{K_{34}T}{V_N}$
$\frac{\left(\frac{MV_N}{GT^2}\right)}{S}$
μ

Fig. 2 Maximum Percentage Headway Error vs. q_1/q_3 for a Headwind Three Times Nominal Velocity: $q_2 = 10q_1$, $q_3 = 10$ and $\tau = .1T$

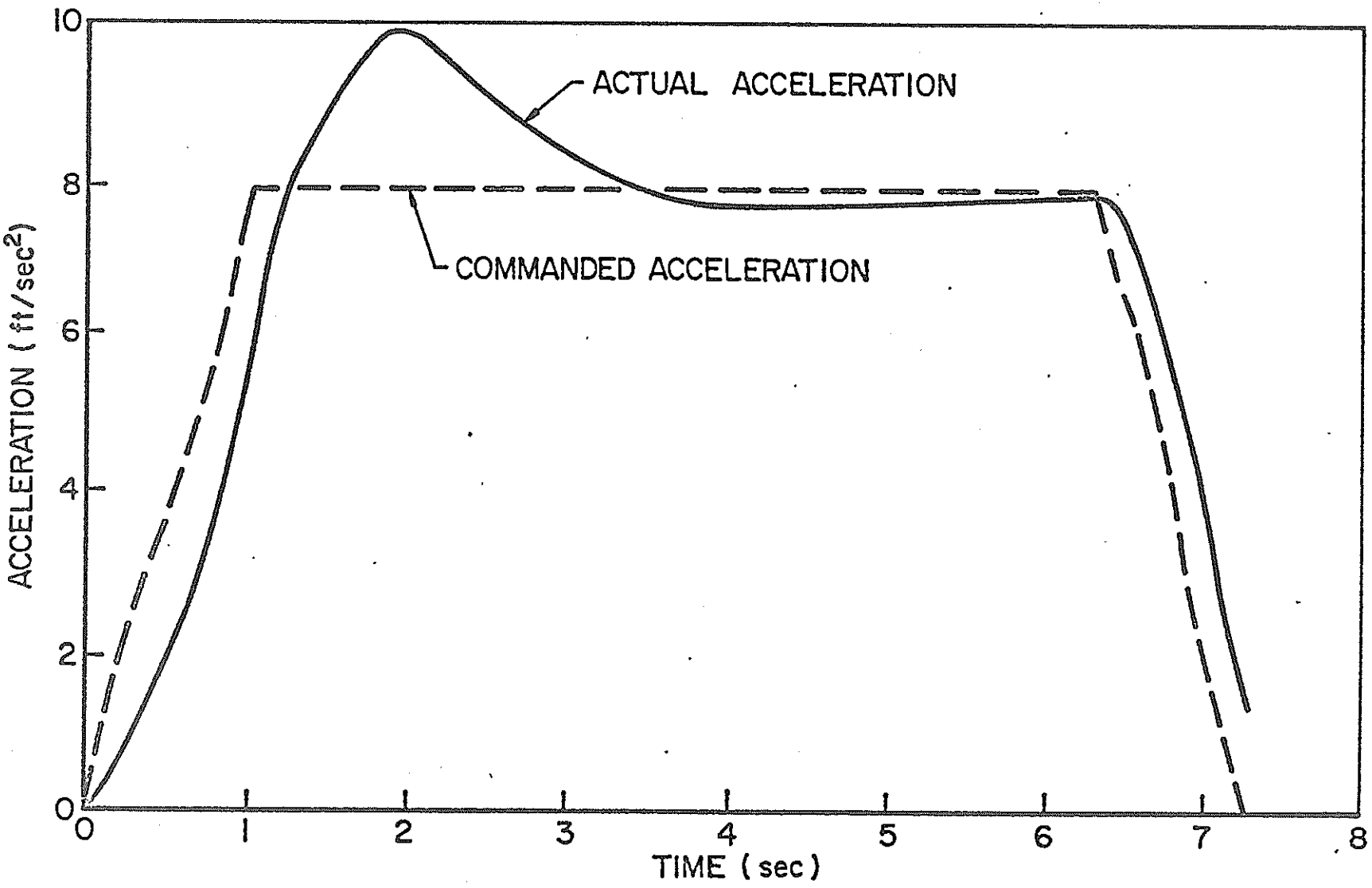

ACTUAL ACCELERATION
COMMANDED ACCELERATION
ACCELERATION (ft/sec2)
TIME (sec)
0
2
4
6
8
10
0
1
2
3
4
5
6
7
8

Fig. 3 Maximum Non-dimensional Acceleration Error vs. q_1/q_3 for a Ten Percent Initial Headway Error: $q_2 = 10q_1$, $q_3 = 10$ and $\tau = .1T$

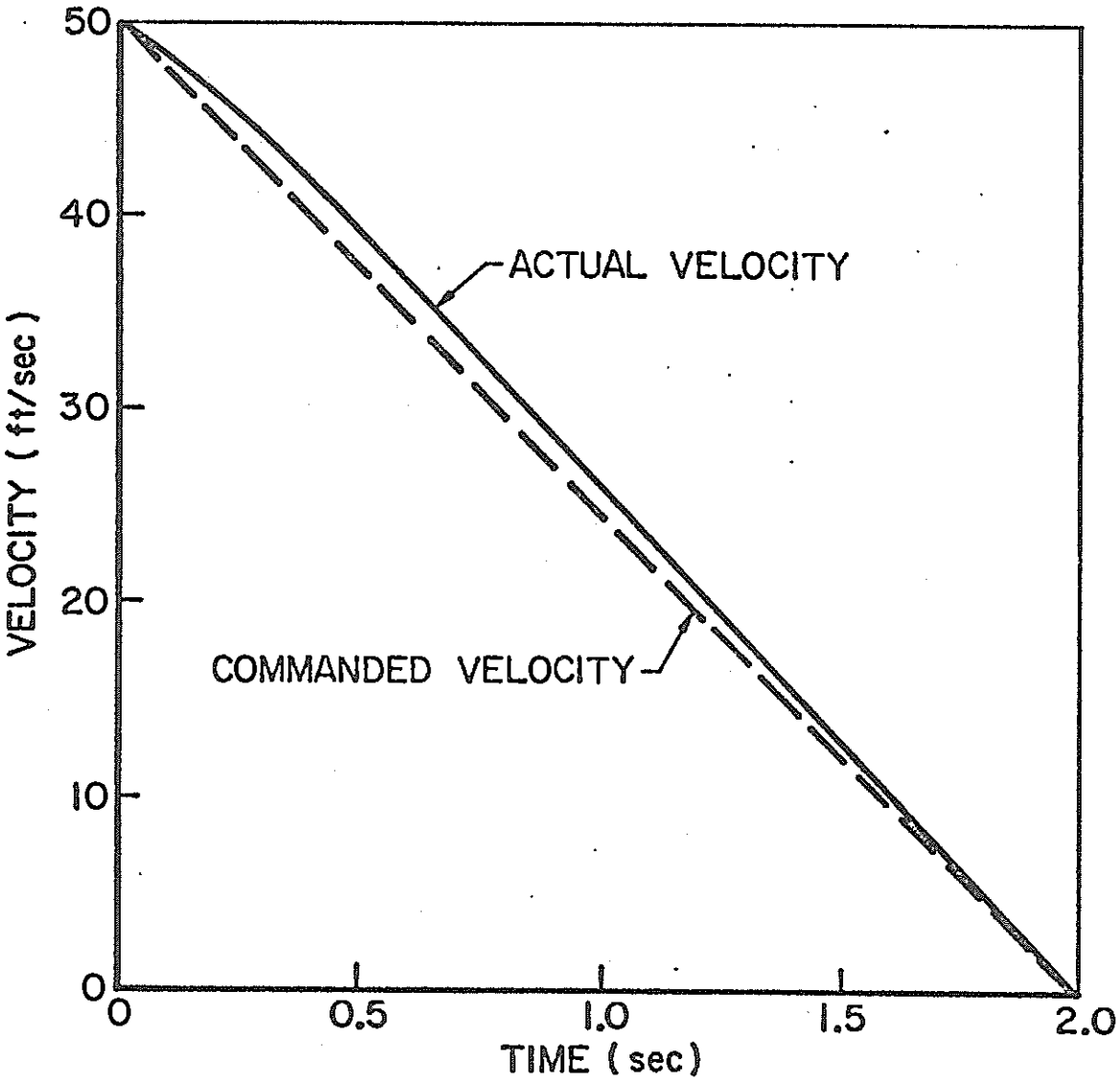

50
40
30
20
10
0
VELOCITY (ft/sec)
ACTUAL VELOCITY
COMMANDED VELOCITY
0
0.5
1.0
1.5
2.0
TIME (sec)

Fig. 4 Maximum Non-dimensional Jerk vs. q_1/q_3 for a Ten Percent Initial Headway Error: $q_2 = 10q_1$, $q_3 = 10$ and $\tau = .1T$

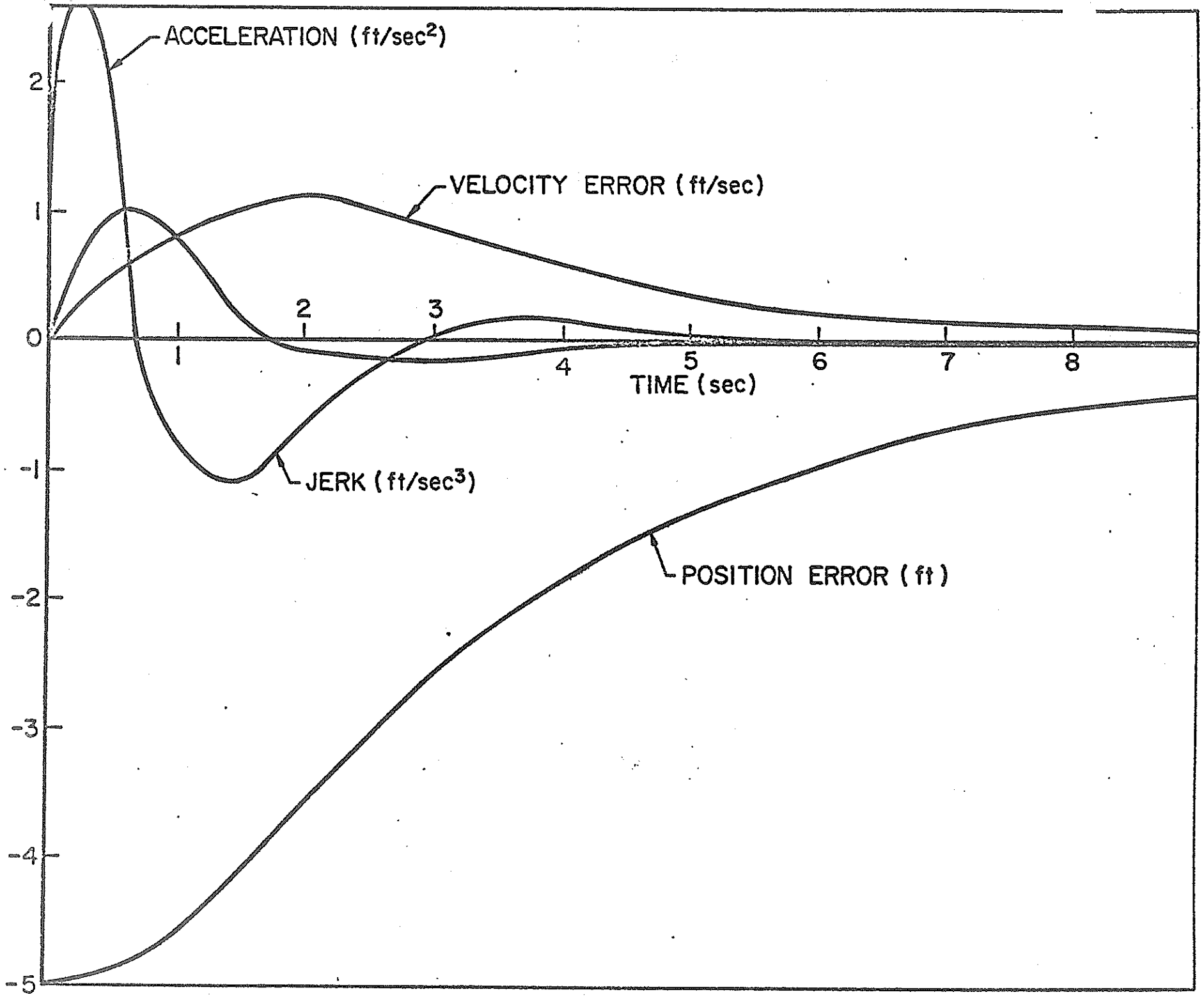

ACCELERATION (ft/sec²)
VELOCITY ERROR (ft/sec)
JERK (ft/sec³)
POSITION ERROR (ft)
TIME (sec)
2
1
0
-1
-2
-3
-4
-5
1
2
3
4
5
6
7
8

Fig. 5 Non-dimensional Time Required to Reduce Headway Error to Ten Percent of Initial Value vs. q_1/q_3: $q_2 = 10q_1$, $q_3 = 10$ and $\tau = .1T$

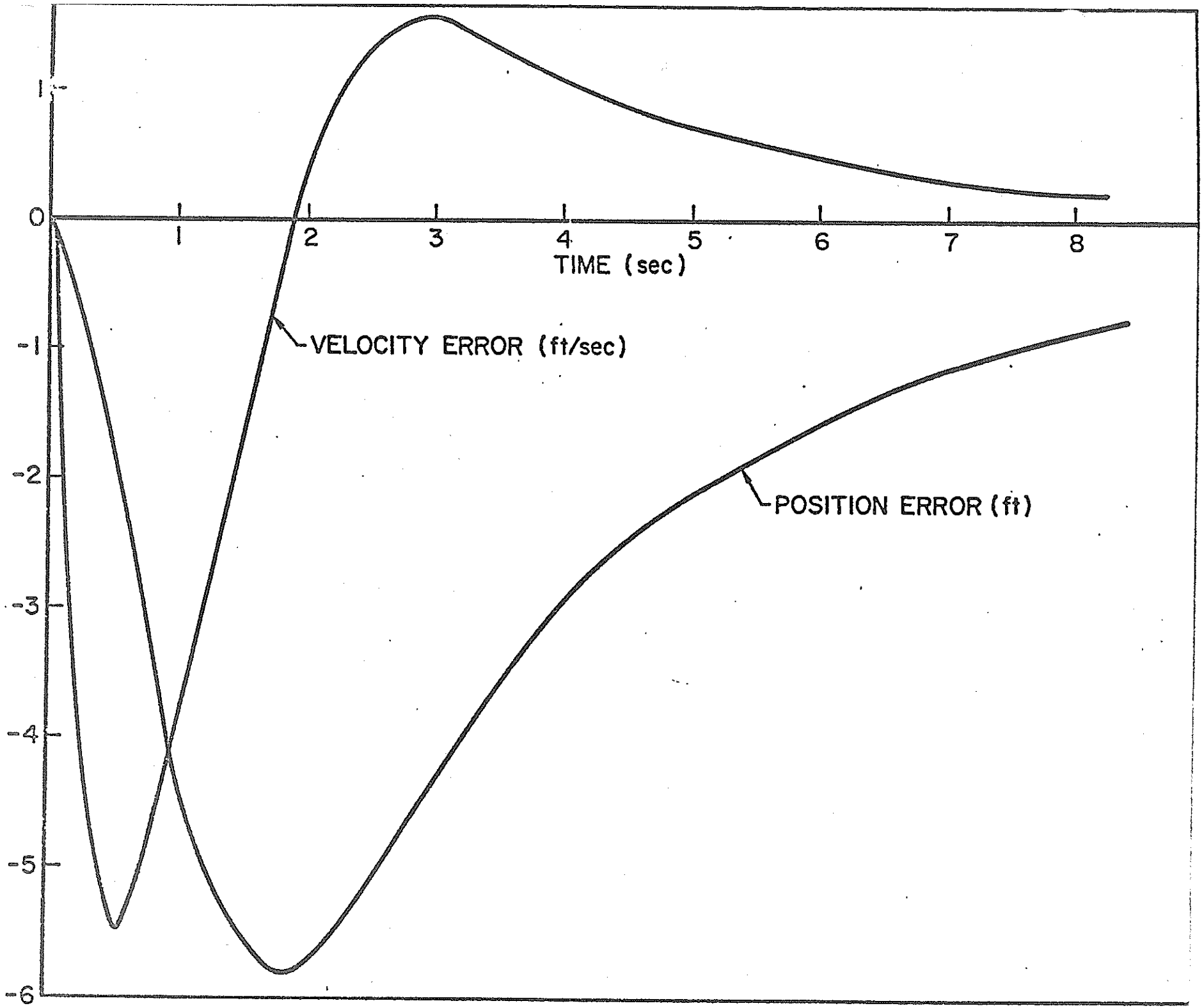

1
0
-1
-2
-3
-4
-5
-6
1
2
3
4
5
6
7
8
TIME (sec)
VELOCITY ERROR (ft/sec)
POSITION ERROR (ft)

Fig. 6 Non-Dimensional Peak Power Required for a Headwind Three Times Vehicle Nominal Velocity vs. q_1/q_3: $q_2 = 10q_1$, $q_3 = 10$ and $\tau = .1T$

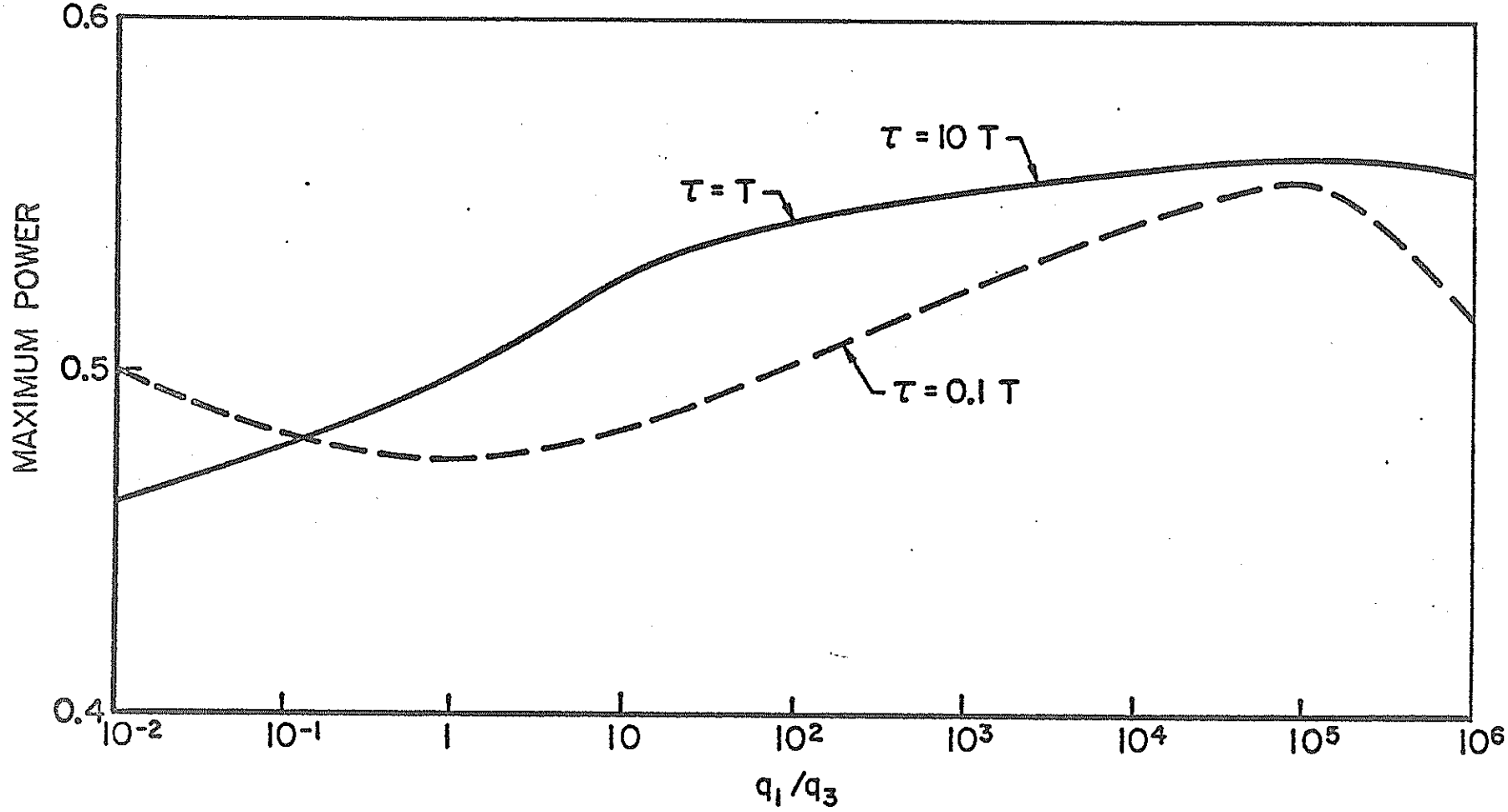

0.6
0.5
0.4
MAXIMUM POWER
τ = 10 T
τ = T
τ = 0.1 T
10-2
10-1
1
10
10²
10³
10⁴
10⁵
10⁶
q1 /q3

Fig. 7 Maximum Percentage Headway Error for Various Propulsion System Time Constants for a Headwind Three Times Vehicle Nominal Velocity: $q_2 = 10q_1$, $q_3 = 10$, and $q_4 = 10$

TIME TO REDUCE HEADWAY ERROR
TO 10% OF INITIAL VALUE
16
12
8
4
0
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6
q_1/q_3
$q_4=10$
$T=0.1\,T$
$T=T$
$T=10\,T$

Fig. 8 Maximum Non-dimensional Acceleration Error for Various Propulsion System Time Constants for Ten Percent Initial Headway Error: $q_2 = 10q_1$, $q_3 = 10$, and $q_4 = 10$

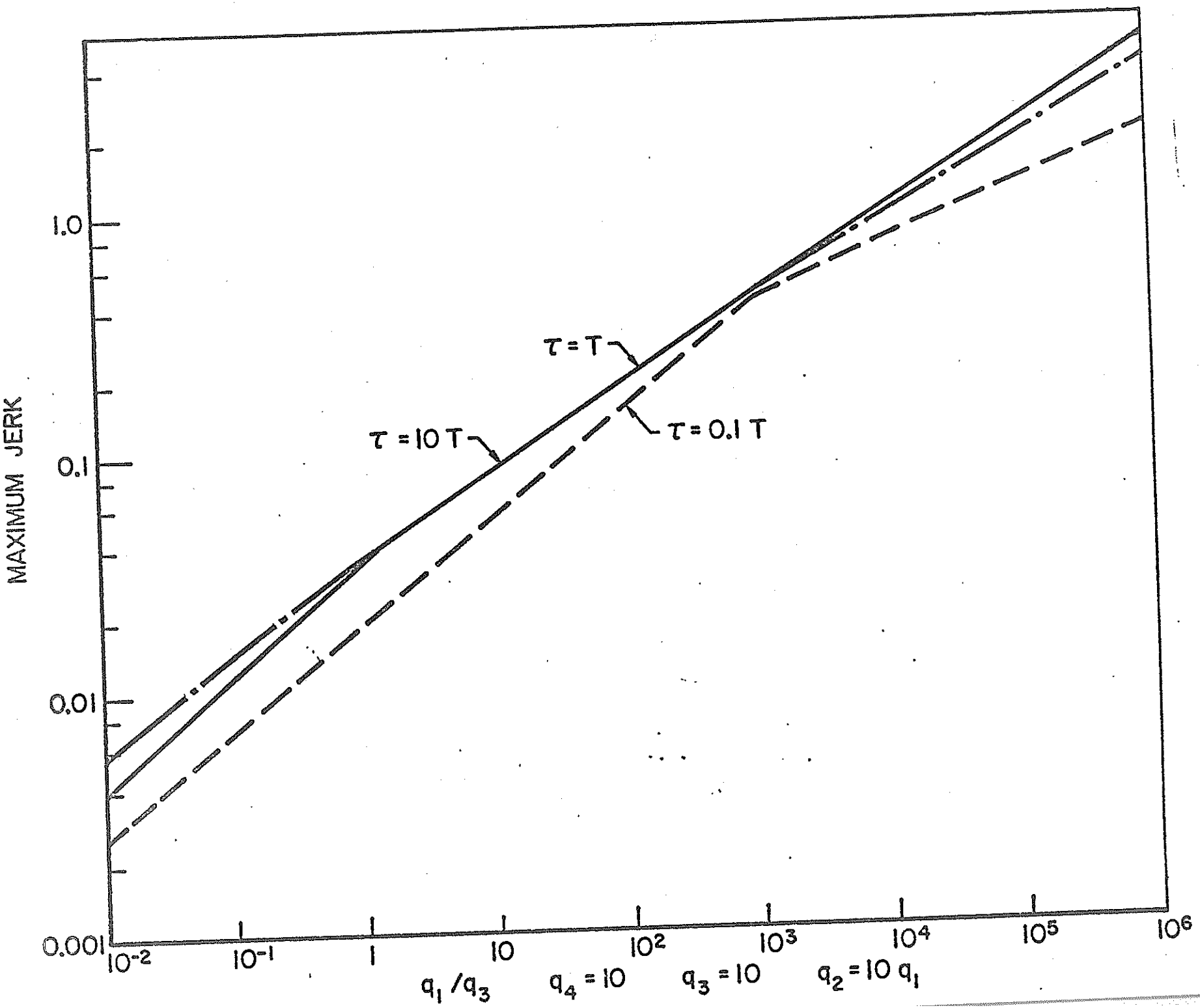

MAXIMUM JERK
1.0
0.1
0.01
0.001
τ = T
τ = 10 T
τ = 0.1 T
10^-2
10^-1
1
10
10^2
10^3
10^4
10^5
10^6
q1 /q3
q4 = 10
q3 = 10
q2 = 10 q1

Fig. 9 Maximum Non-dimensional Jerk for Various Propulsion System Time Constants for a Ten Percent Initial Headway Error: $q_2 = 10q_1$, $q_3 = 10$, and $q_4 = 10$

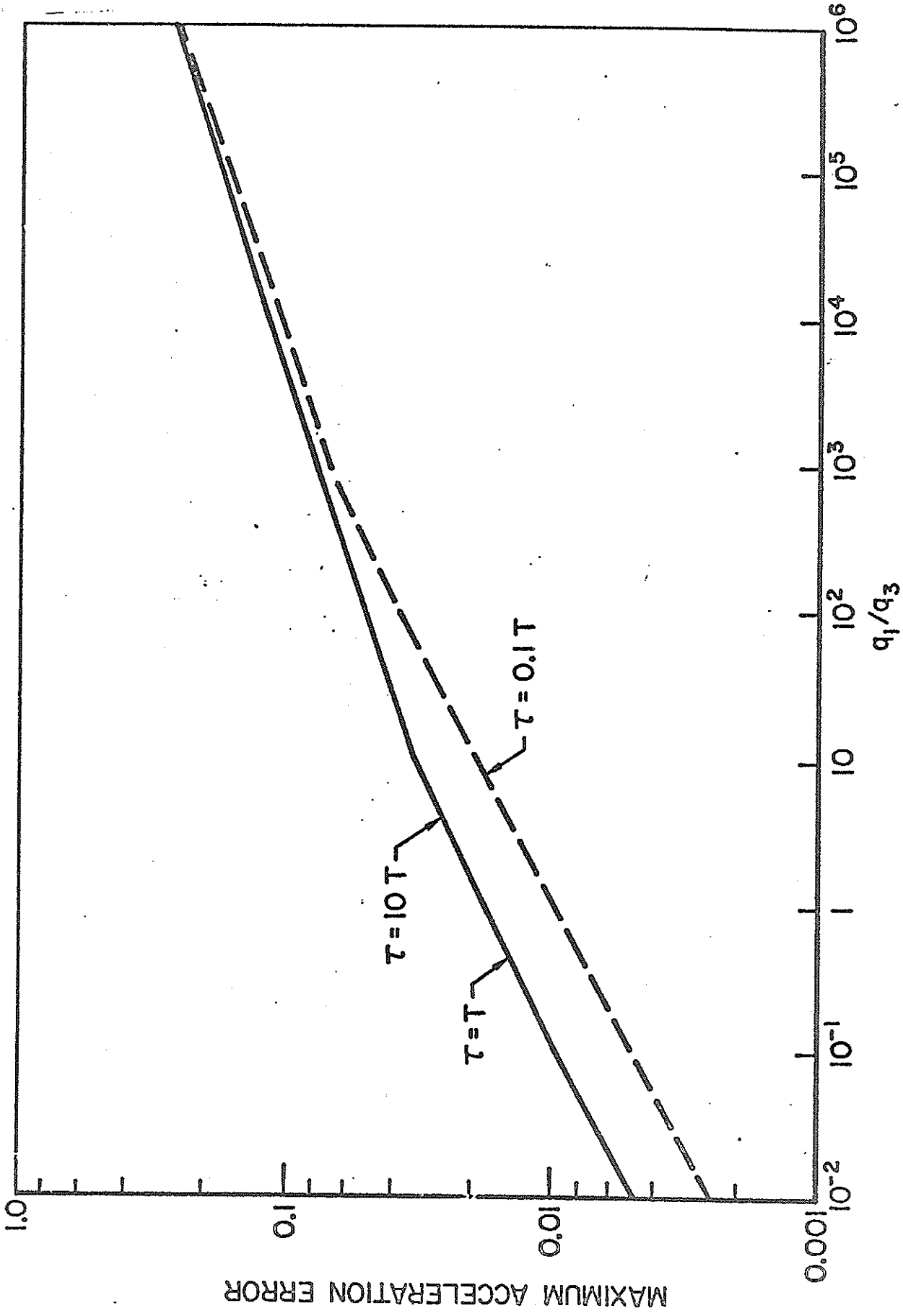

MAXIMUM ACCELERATION ERROR
1.0
0.1
0.01
0.001
10^{-2}
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6
q_1/q_3
$\tau = 0.1\,T$
$\tau = 10\,T$
$\tau = T$

Fig. 10 Non-dimensional Time Required to Reduce Headway Error to Ten Percent of Initial Value for Various Propulsion System Time Constants: $q_2 = 10q_1$, $q_3 = 10$, and $q_4 = 10$

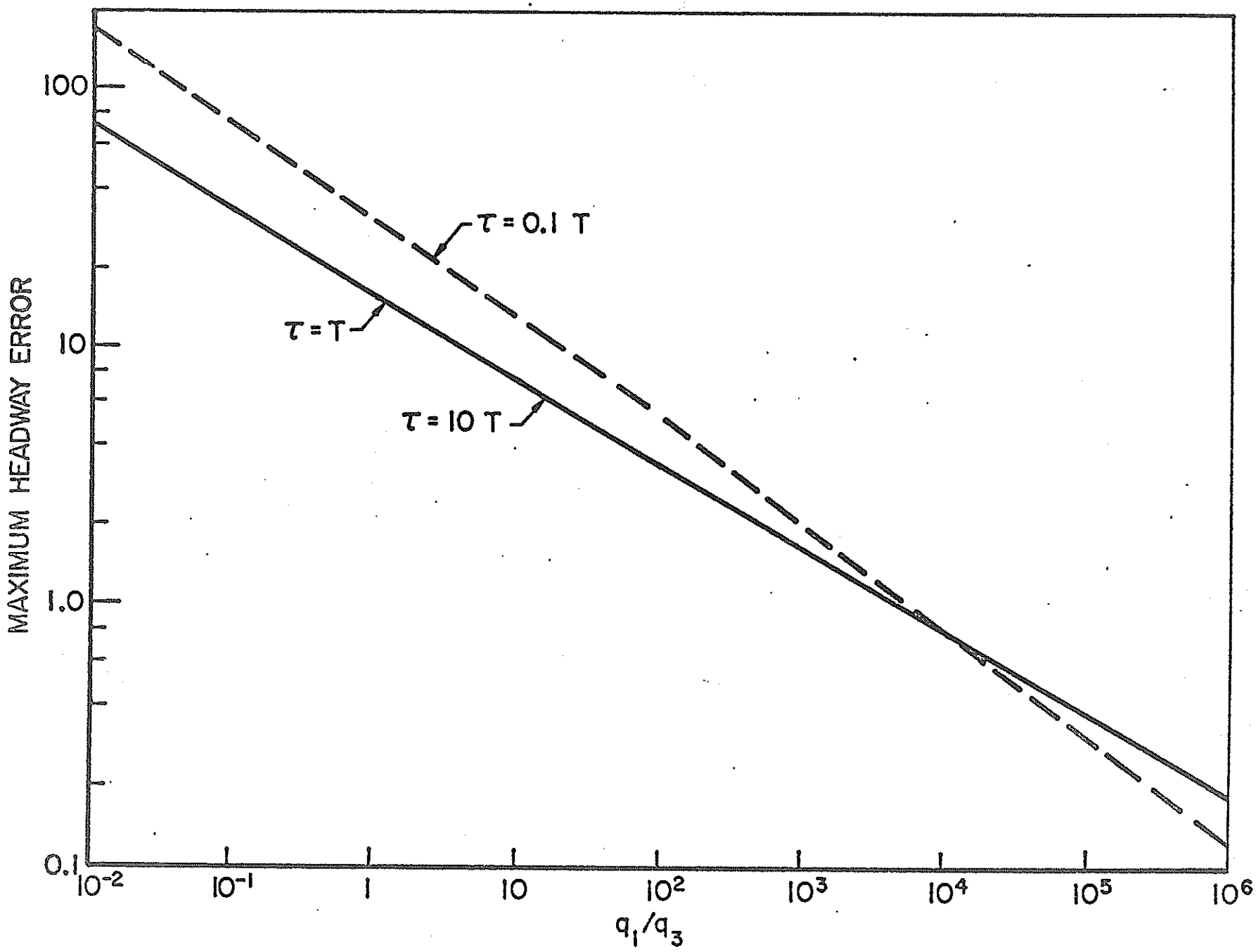

MAXIMUM HEADWAY ERROR
100
10
1.0
0.1
τ = 0.1 T
τ = T
τ = 10 T
10^-2
10^-1
1
10
10^2
10^3
10^4
10^5
10^6
q1/q3

Fig. 11 Non-dimensional Peak Power Required for a Headwind Three Times Vehicle Nominal Velocity for Various Propulsion System Time Constants: $q_2 = 10q_1$, $q_3 = 10$ and $q_4 = 10$

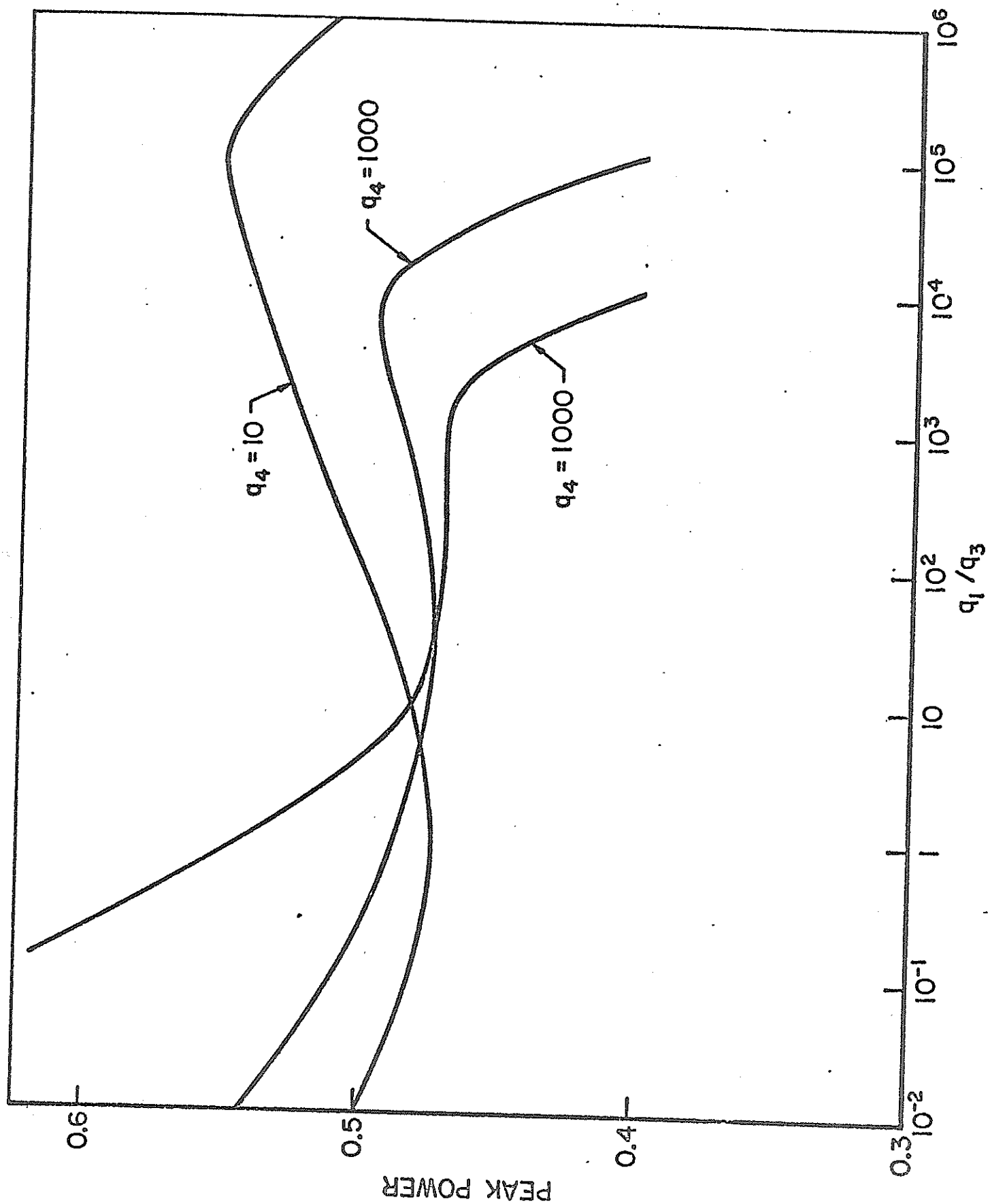

PEAK POWER
0.6
0.5
0.4
0.3
10^{-2}
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6
q_1/q_3
$q_4 = 10$
$q_4 = 1000$
$q_4 = 1000$

Fig. 12 Response of Vehicle to a Suddenly Applied Headwind of 150 ft/sec

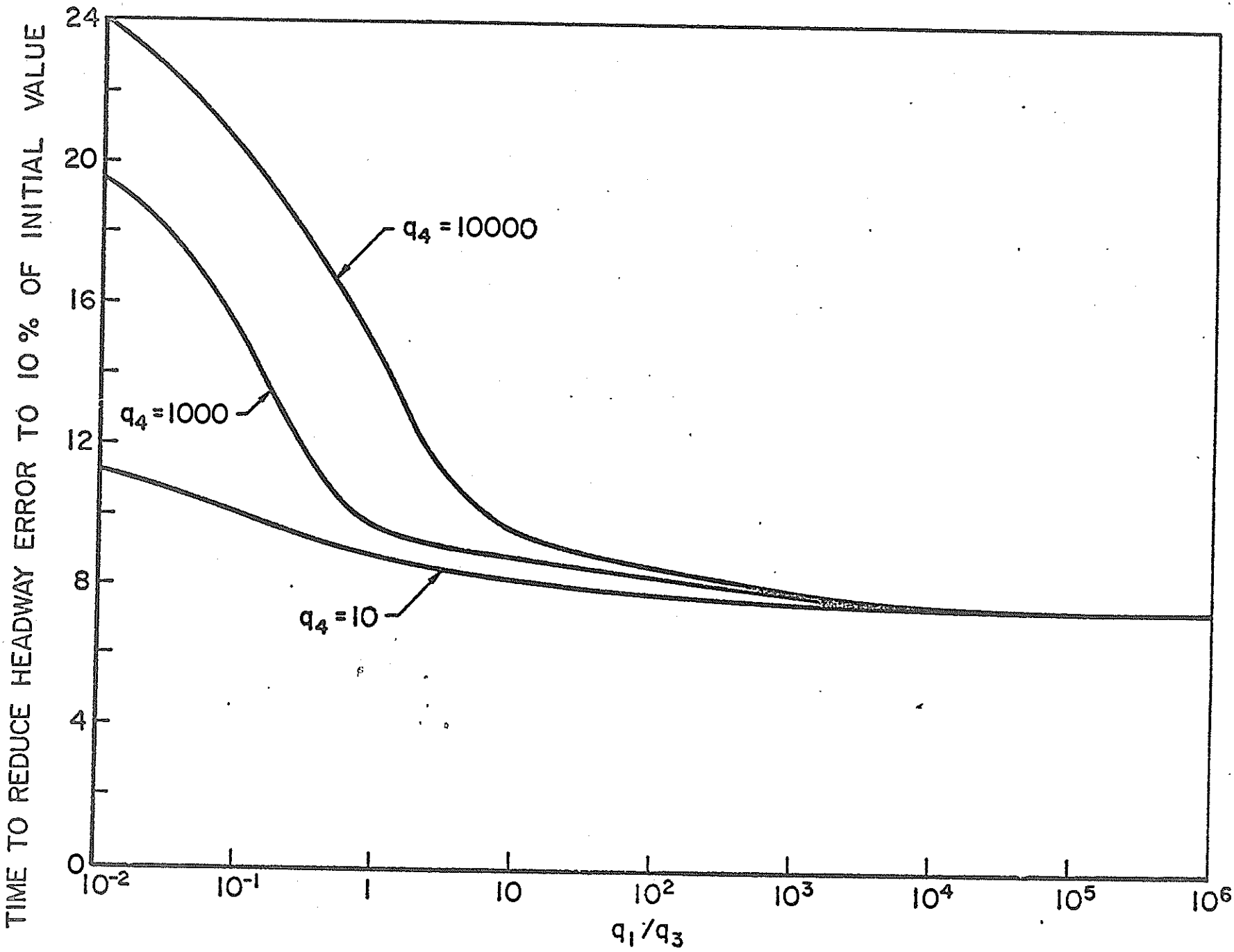

TIME TO REDUCE HEADWAY ERROR TO 10 % OF INITIAL VALUE
24
20
16
12
8
4
0
$q_4 = 10000$
$q_4 = 1000$
$q_4 = 10$
10^{-2}
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6
q_1/q_3

Fig. 13 Response of Vehicle to a 5 ft (10%) Initial Headway Error

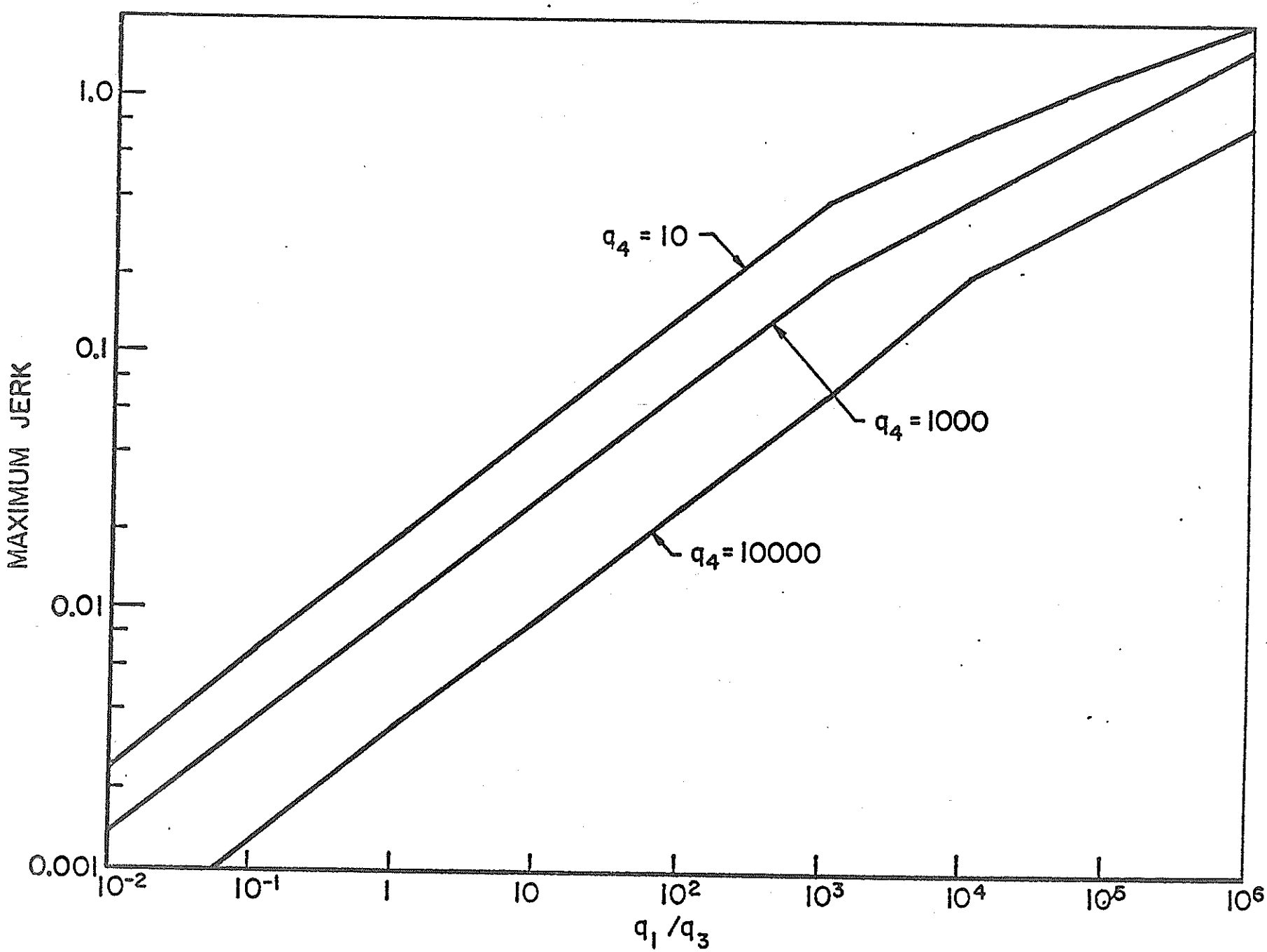

MAXIMUM JERK
1.0
0.1
0.01
0.001
$q_4 = 10$
$q_4 = 1000$
$q_4 = 10000$
10^{-2}
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6
q_1 / q_3

Fig. 14 Vehicle Velocity During Emergency Stopping (Emergency Deceleration 25 ft/sec^2)

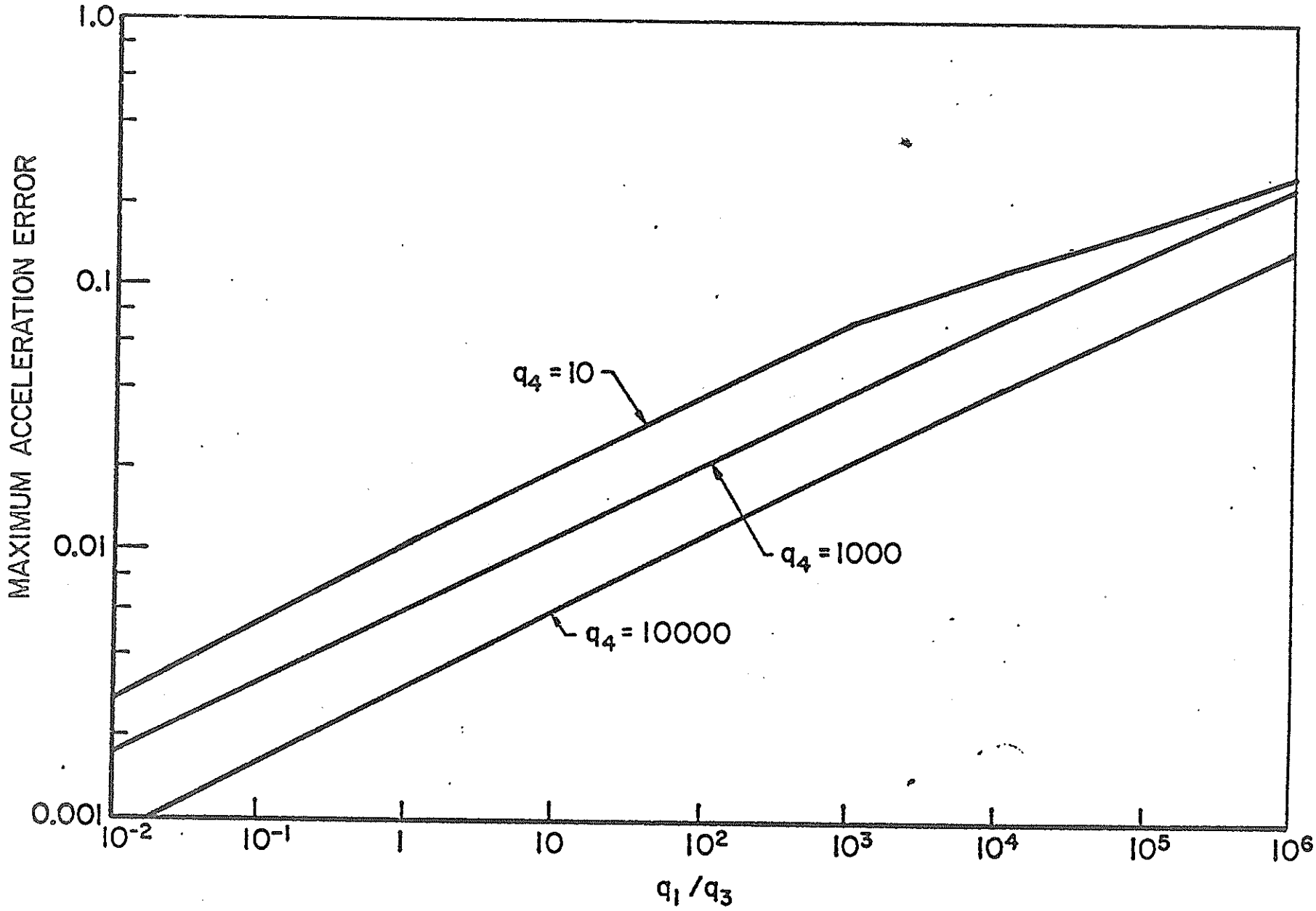

1.0
0.1
0.01
0.001
MAXIMUM ACCELERATION ERROR
$q_4 = 10$
$q_4 = 1000$
$q_4 = 10000$
10^{-2}
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6
q_1 / q_3

Fig. 15 Vehicle Acceleration During Merging

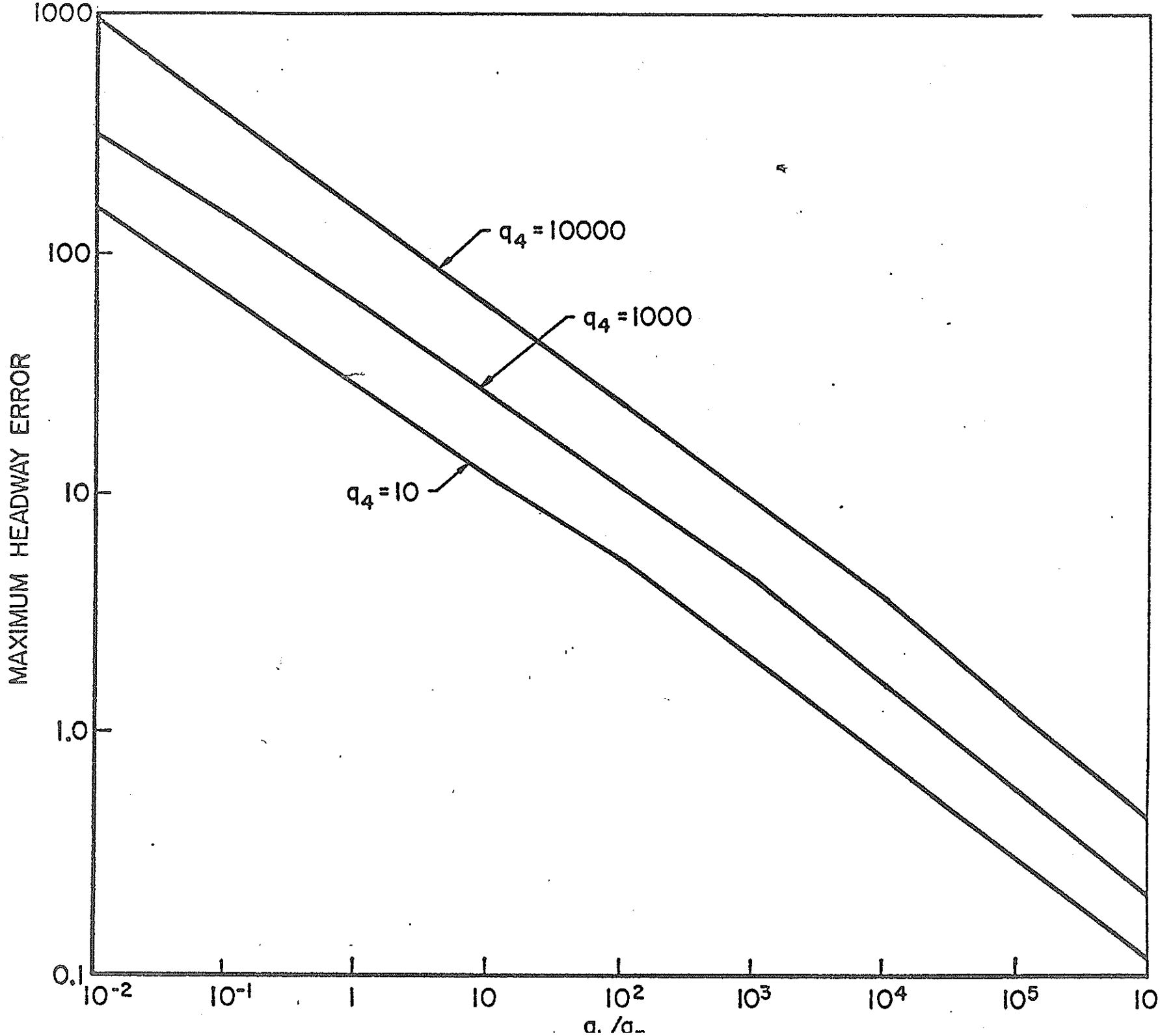
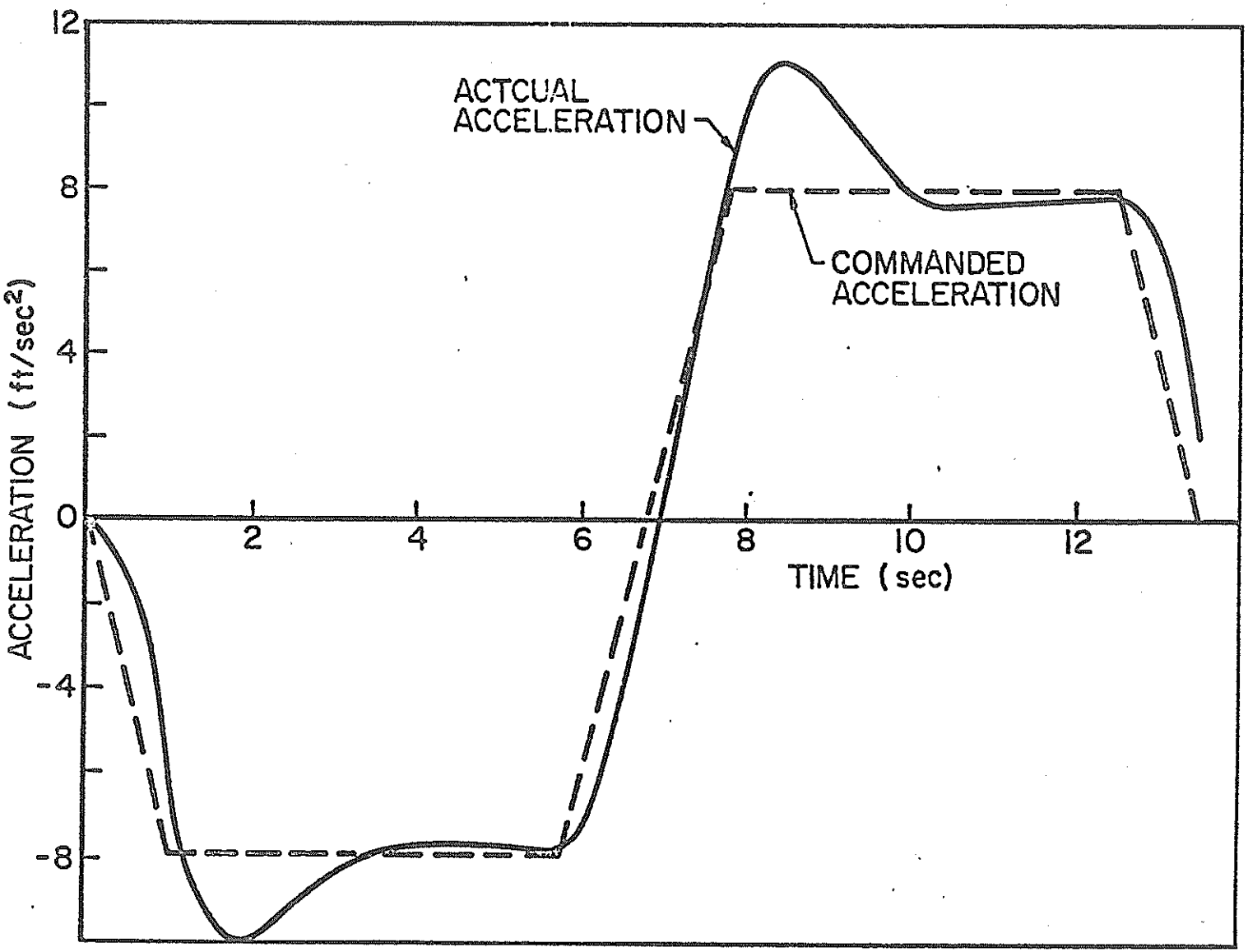

1000
100
10
1.0
0.1
MAXIMUM HEADWAY ERROR
$q_4 = 10000$
$q_4 = 1000$
$q_4 = 10$
10^{-2}
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6
q_1/q_4

Fig. 16 Vehicle Acceleration During Maneuvering (Slipping One Headway Slot)


12
8
4
0
-4
-8
2
4
6
8
10
12
ACCELERATION (ft/sec²)
TIME (sec)
ACTCUAL
ACCELERATION
COMMANDED
ACCELERATION

Use of State Observers in the Optimal Feedback Control of Automated Transit Vehicles

W. L. Garrard and A. L. Kornhauser

Abstract

The theory of optimal control and the theory of observers is applied to the design of feedback systems for longitudinal control of vehicles in automated, high-capacity transit systems. The resulting controllers require only measurement of position and velocity errors and excellent dynamic response is achieved for mainline operation, for merging and demerging from stations, for manuevering at intersections and for emergency stopping.

Use of State Observers in the Optimal Feedback Control of Automated Transit Vehicles

W. L. Garrard* and A. L. Kornhauser**

1. INTRODUCTION

Almost all new urban transportation systems involve the use of automatically controlled vehicles. In many systems, such as personal rapid transit (PRT), dual-mode, and automated freeways, the number of passengers per vehicle is small and short headways are necessary for high capacity. This requires a versatile and efficient control system which must maintain the proper spacing between vehicles without causing passenger discomfort. The control system must be reasonably economical to implement, adaptable to merging and demerging from off-line stations and maneuvering at interchanges, simple enough to insure reliability, and suitable for use in emergency situations[1].

In a previous study optimization theory was applied to the design of feedback control systems for high capacity transit systems such as PRT or dual-mode[2]. The resulting optimal controllers were shown to keep headway and velocity errors small

---

* Department of Aerospace Engineering and Mechanics and Center for Control Science, University of Minnesota, Minneapolis, Minnesota.

** Department of Civil Engineering, Princeton University, Princeton, New Jersey. (Member ASME).

without causing passenger discomfort. Computer simulations indicated that excellent dynamic response could be achieved for normal mainline operation with headwinds three times the nominal mainline vehicle velocity, for merging and demerging, for maneuvering, and for emergency stopping.

The appropriate state variables for the longitudinal control problem were determined to be the position error, velocity error, acceleration error, and rate of change of propulsive force. Implementation of the optimal controller necessitated measurement of all of these variables. In actual practice, it is inconvenient and expensive to measure accurately the acceleration error and rate of change of propulsive force. In this paper, the theory of observers is applied to the estimation of all of the state variables from measurements of only position and velocity errors[3] The dynamic response resulting from use of a controller in which acceleration error and rate of changes of propulsive force are estimated is shown to compare favorably with that achieved when all state variables are measured. Also the effects on system performance of sampling of position error are considered.

2. VEHICLE DYNAMICS AND CONTROL SYSTEMS MODEL

2.1 System Operating Philosophy

This study focuses on the design of longitudinal controllers for vehicles operating transit systems which function in a synchronous or quasi-synchronous manner. In both synchronous and quasi-synchronous operation, vehicles follow hypothetical slots moving along the guideway at nominal line velocity. The slot lengths are uniform and are equal to the length of the vehicle plus the

minimum allowable nose to tail separation between adjacent vehicles. Vehicles do not communicate directly with one another and the position of a vehicle is determined with respect to the guideway rather than with respect to other vehicles in the system [4-7]. Vehicles may be maneuvered from one slot to another in order to resolve conflicts at intersections and merge points. The vehicles communicate with wayside computers which manage traffic on the mainline and at stations and intersections. These computers command the maneuvers necessary to avoid conflicts and also institute emergency procedures in case of a vehicle failure.

The longitudinal control system must be capable of holding a vehicle within its alloted slot under the action of headwinds and other disturbances without causing passenger discomfort. In addition, it must be capable of closely following one of a set of acceleration-deceleration profiles as commanded by the wayside computers during merging and demerging, maneuvering to avoid conflicts at intersections or to push a failing vehicle, or stopping for emergencies. The remainder of this paper is concerned with the design of such a control system.

2.2 Equation of Motion

The differential equation describing the longitudinal motion of the transit vehicle is

$$M\frac{dV}{dt} = -F_D(V,V_W) + F - Mg \sin\theta - F_M \qquad (1)$$

where:

M = mass of the vehicle

V = velocity of the vehicle

V_W = velocity of the wind (positive for a head wind)

F = propulsive force

g = gravitational acceleration

θ = slope of the guideway

F_D = aerodynamic drag

F_M = mechanical resistance

The aerodynamic drag is

$$F_D = C_D(V + V_W)^2 \tag{2}$$

where C_D is a drag coefficient. Furthermore, the propulsive force is assumed to be governed by

$$\frac{dF}{dt} = -(\frac{1}{\tau})F + Gi \tag{3}$$

where:

τ = time constant of the propulsion system

i = control input to the propulsion system

G = gain constant of the propulsion system.

That is, the propulsion system as modeled as a first-order lag. The error, e, is defined as

$$e = X - X_c \tag{4}$$

where

X = the actual position of the vehicle

X_c = the desired or command position of the vehicle.

Since $V = \dot{X}$, (1) can be re-written in terms of the error as

$$\frac{d^2e}{dt^2} = -\frac{C_D}{M}\left(\frac{de}{dt} + \frac{dXc}{dt} + V_W\right)^2 + \frac{F}{M} - g \sin \theta - \frac{d^2X_c}{dt^2} - \frac{F_M}{M} . \quad (5)$$

2.3 Nondimensional Formulation

For purposes of generality, the system equations will be non-dimensionalized. The following non-dimensional variables will be used:

$v = \frac{T^2 Gi}{MV_N}$, non-dimensional control input to the propulsion system

$w = \frac{Vw}{V_N}$ non-dimensional headwind velocity

$y = \frac{e}{H}$, non-dimensional error

$y_c = \frac{X_c}{H}$, non-dimensional command position

$\sigma = \frac{t}{T}$, non-dimensional time

$f = \frac{TF}{MV_N}$, non-dimensional propulsive force

$\cdot = \frac{d}{d\sigma}$, derivative with respect to non-dimensional time

where:

V_N = nominal velocity of the vehicles on the main guideway

H = nominal nose to nose distance between vehicles on the main guideway

T = nominal time headway between vehicles, $T = \frac{H}{V_N}$.

The resulting system equations are

$$\ddot{y} = -\left(\frac{C_D}{M}\right) H(\dot{y} + \dot{y}_c + w)^2 + f - \frac{Tg \sin\theta}{V_N} - \ddot{y}_c - \frac{F_M T}{MV_N} \qquad (6)$$

$$\dot{f} = -\frac{T}{\tau} f + v \qquad (7)$$

During normal mainline operation the vehicle will operate at near nominal velocity, and linearization of (6) and (7) is legitimate. Furthermore $\ddot{y}_c$, the commanded acceleration, will be zero, and $\dot{y}_c$, the commanded velocity, will be unity. The resulting linearized equations of motion are

$$\begin{aligned} \ddot{y} &= -\left(\frac{2C_D}{M}\right) H (1 + w)\dot{y} + f - d \\ \dot{f} &= -\frac{T}{\tau} f + v \end{aligned} \qquad (8)$$

$d = \frac{Tg \sin\theta}{V_N} + \frac{F_M T}{V_N M} + \left(\frac{C_D H}{M}\right)(w + 1)^2$, the non-dimensional disturbance force.

The velocity of the headwind is not known a priori; however, its average value will in general be zero. Thus the coefficient of $\dot{y}$ in (8) will be approximated by its average value $\frac{2C_D H}{M}$. It has been shown that a feedback controller designed on the basis of the above assumptions provides excellent dynamic response for mainline operation with headwinds three times the nominal vehicle velocity; for merging and demerging from off-line stations; for maneuvering at interchanges; and for emergency control [2].

3. SYNTHESIS OF THE OPTIMAL FEEDBACK CONTROL SYSTEM

3.1 State Variables

The vehicle and propulsion system have been modeled by a set of linear differential equations with constant coefficients. If the state variables are selected properly, it is possible to use optimization theory to design a feedback control system which will keep headway and velocity errors small without causing passenger discomfort. As shown previously [2], the appropriate state variables for this problem are headway error, velocity error, acceleration error, and rate of change of propulsive force. Headway and velocity error are obvious choices for state variables; however, acceleration error and rate of change of propulsive force are less obvious candidates. The reasons for selecting these two quantities as state variables will be discussed below.

The state variables are $x_1 = y$, the non-dimensional headway error; $x_2 = \dot{y}$, the non-dimensional velocity error; $x_3 = \ddot{y}$, the non-dimensional acceleration error; and $x_4 = \dot{f}$, the rate of change of the non-dimensional propulsive force. The control variable is $u = \dot{v}$, the rate of change of the non-dimensional input to the propulsion system. For purposes of control system design, the non-dimensional disturbance force, d, is assumed constant (the headwind is constant or the vehicle is ascending or descending a constant slope). Using these definitions and assumptions, the equations of motion of the vehicle can be written in vector-matrix form as

$$\dot{\underline{x}} = A\underline{x} + \underline{b}u \qquad (9)$$

where

$$\underline{x}^T = [x_1 \; x_2 \; x_3 \; x_4],$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{2C_D H}{M} & 1 \\ 0 & 0 & 0 & -\frac{T}{\tau} \end{bmatrix},$$

and

$$\underline{b}^T = [0 \quad 0 \quad 0 \quad 1].$$

The control will be selected in such a manner as to drive the state variables to zero. The state variables x_1 and x_2 represent the position and velocity errors and the necessity of driving these quantities to zero is clear. The state variables x_3 and x_4 represent the acceleration error and the rate of change of propulsive force. The acceleration error must be zero if the position and velocity errors are to remain zero; furthermore, in order to achieve zero acceleration error, the propulsive force must equal the disturbance force. Since the disturbance force is assumed to be constant, the propulsive force must also approach a constant value, and the derivative of the propulsive force, x_4, must approach zero.

3.2 The Performance Index

The error equations have been formulated in the standard notation of optimal control theory; however, in order to apply

this theory a mathematical criterion for the measurement of system performance is necessary. A quadratic performance index is proposed. This index is

$$J = \frac{1}{2} \int_0^\infty (q_1 x_1^2 + q_2 x_2^2 + q_3 x_3^2 + q_4 x_4^2 + r u^2)\, dt. \qquad (10)$$

The optimal feedback control problem is to determine the control, u, as a function of the state variables in such a manner as to minimize J. It is well known that the control which minimizes J drives the state variables to zero [8].

The performance index J is the integral of the weighted sum of the position error, velocity error, acceleration error, derivative of the propulsive force, and the control effort. Since the drag coefficient is small, the derivative of the propulsive force is very nearly equal to the jerk. The performance index penalizes large position and velocity errors and the control which minimizes J should result in a system in which these errors are kept small. The performance index also penalizes large acceleration errors and jerks. These two variables affect passenger comfort and the control which minimizes J should also result in a system in which passenger discomfort is minimized. The rate of change of the control input to the propulsion system, u, must be included in the performance index in order to obtain the optimal control in feedback form.

It is of course possible to formulate many other performance indices which include system error and passenger comfort;

however, use of a quadratic performance index as given in (10) permits determination of the optimal control in feedback form. This is one of the few classes of optimization problems in which the optimal feedback control can be found [8]. In addition the optimal feedback control is linear with constant gains. Such a controller is easy to implement.

The feedback control which minimizes J is

$$u = -r^{-1}\,\underline{b}^T K \underline{x} \tag{11}$$

where K, the optimal gain matrix, is the symmetric, positive-definite solution of the matrix Ricatti equation

$$KA + A^T K - K\underline{b}^T r^{-1}\underline{b}K + Q = 0 \tag{12}$$

where

$$Q = \begin{bmatrix} q_1 & 0 & 0 & 0 \\ 0 & q_2 & 0 & 0 \\ 0 & 0 & q_3 & 0 \\ 0 & 0 & 0 & q_4 \end{bmatrix} .$$

Iterative methods for the solution of (12) allow rapid determination of the optimal gain matrix by use of a high-speed digital computer [9].

From (12) the optimal control in terms of the actual error variables is

$$u = -\frac{K_{14}}{H}e - \frac{K_{24}}{V_N}\dot{e} - \frac{K_{34}T\ddot{e}}{V_N} - \frac{K_{44}T^2}{MV_N}\dot{F} \tag{13}$$

where the K_{ij}'s are the elements of the optimal gain matrix K. The optimal control (13) is similar to the control derived by Whitney and Tomisuka [6] using classical techniques and by Wilkie [7] using optimization theory. However, in neither of these studies was the dynamics of propulsion system considered.

3.3 Determination of Weighting Matrix Q.

The weighting factors q_1, q_2, q_3, q_4 and r affect the values of the elements of K, the gain matrix.* The values of the elements of the gain matrix in turn affect the dynamic response of the system. The relationship between the weighting factors and the dynamic response of the vehicle cannot be analytically determined. If the weighting factors on headway and velocity error are chosen to be large relative to the weighting factors on the acceleration errors and jerk (rate of change of propulsive force) a system which zeros errors rapidly but which gives an uncomfortable ride will result. On the other hand, if the weighting factors on acceleration error and jerk are chosen to be large relative to the weighting factors on headway and velocity error, the ride will be comfortable but the system will be rather sluggish in reducing the headway error to zero. Thus the designer must determine how the weighting factors affect the dynamic response in order to obtain the proper trade-off between ride quality and adequate control of headway error. The relationship between the

---

* It can be shown that $K_{14} = \sqrt{q_1}$; the other gains must be determined numerically, however.

weighting factors and the dynamic response of the optimally controlled vehicle was studied in detail by Garrard and Kornhauser [2] for several values of the propulsion system time constant.

4. IMPLEMENTATION OF THE OPTIMAL CONTROLLER

4.1 Development of the Observer

Implementation of the optimal feedback controller necessitates measurement of all state variables - the position error, velocity error, acceleration error, and rate of change of propulsive force. In actual practice, it is inconvenient and expensive to measure accurately the acceleration error and rate of change of propulsive force. However it is possible, to estimate or reconstruct the values of these variables from measurements of the position and velocity errors. This estimation can be accurately performed by using the theory of observers [3].

For design purposes, the vehicle and propulsion system is modeled by (9), and the optimal control is

$$u = -\underline{k}^T \underline{x} \qquad (14)$$

Where from (11), $\underline{k}^T = r^{-1}\underline{b}^T K$. The measurable output of the system is defined as a vector, $\underline{y}$, of dimensionality $p \leq 4$. This vector is given by

$$\underline{y} = C\underline{x} \qquad (15)$$

where C is a matrix of dimensionality p x 4.

In order to implement the optimal control, $\underline{x}$ must be reconstructed from $\underline{y}$. This is possible if and only if the system is observable [3]. It can easily be shown that the position error of the vehicle must be measured in order to guarantee observability. The velocity error, acceleration error, and derivative of propulsive force can be reconstructed from the position error alone. However, it is relatively simple to measure velocity error; therefore, the matrix C is assumed to be

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{16}$$

That is, both position and velocity errors are measured.

The theory of observers can be used to synthesize a network the output of which is a suitable approximation of the state of the vehicle. An observer is a dynamical system which operates on the output of another dynamical system in order to provide an estimate of the state of that system [3]. The estimate of the state of the original system is denoted as $\hat{\underline{x}}$, and the state of the observer is defined as $\underline{z}$ where

$$\underline{z} = T\hat{\underline{x}} \tag{17}$$

and

$$\dot{\underline{z}} = F\underline{z} + G\underline{y} + Tb\underline{u} \tag{18}$$

with

$$TA - FT = GC \tag{19}$$

In the longitudinal control problem as formulated in this study, $\underline{z}$ is of dimensionality two, and

$$\hat{\underline{x}} = \begin{bmatrix} \underline{C} \\ \underline{T} \end{bmatrix}^{-1} \begin{bmatrix} \underline{y} \\ \underline{z} \end{bmatrix} \tag{20}$$

The matrices F and G are 2 x 2 and the matrix T is 2 x 4. Thus sixteen matrix elements have been introduced. Matrix equation (19) represents eight algebraic equations relating the elements of F, G and T, thus eight of the matrix elements can be chosen arbitrarily. These elements will be selected in such a manner as to result in a simple, efficient observer.

First the matrix F is selected as

$$F = \begin{bmatrix} \lambda_1 & \beta \\ 0 & \lambda_2 \end{bmatrix} \tag{21}$$

The reasons for selecting F in this form will be given later. Substitution of (21) into (19) yields

$$\lambda_1 t_{11} - \beta t_{21} = g_{11} \tag{22a}$$

$$-\lambda_2 t_{21} = g_{21} \tag{22b}$$

If $g_{11} = g_{21} = 0$, $t_{11} = t_{21} = 0$ and the structure of the observer is simplified with no apparent degradation of performance. After making this simplification, the remaining equations resulting from the substitution of (21) into (19) are

$$-\lambda_1 t_{12} - \beta t_{22} = g_{12} \tag{23a}$$

$$-\lambda_2 t_{22} = g_{22} \tag{23b}$$

$$t_{12} + (\ell-\lambda_1)t_{13} - \beta t_{23} = 0 \tag{23c}$$

$$t_{22} + (\ell-\lambda_2)t_{23} = 0 \tag{23d}$$

$$t_{13} + (m-\lambda_1)t_{12} - \beta t_{24} = 0 \tag{23e}$$

$$t_{24} + (m-\lambda_2)t_{24} = 0 \tag{23f}$$

where $\ell = \frac{-2C_D H}{M}$ and $m = \frac{-\tau}{T}$.

Now

$$\begin{bmatrix} C \\ \hline T \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{t_{22}t_{14}-t_{12}t_{24}}{\Delta} & \frac{t_{24}}{\Delta} & -\frac{t_{14}}{\Delta} \\ 0 & \frac{t_{12}t_{23}-t_{22}t_{13}}{\Delta} & -\frac{t_{23}}{\Delta} & \frac{t_{13}}{\Delta} \end{bmatrix} \tag{24}$$

where $\Delta = \beta t_{24}^2$. Obviously $\beta \neq 0$. It can also be shown that $t_{24} = 0$ if $g_{22} = 0$; therefore, $g_{22} \neq 0$. A block diagram of the observer, vehicle and propulsion system is shown in Fig. 1.

It can be shown that an observer does not change the eigenvalues of the original system but simply adjoins its eigenvalues to those of the original system [3]. Thus if the original system

is stable and the eigenvalues of the observer are chosen to have negative real parts, the resulting system will also be stable. If the eigenvalues of the observer are chosen to be large, the estimate of the state will approach the actual value of the state extremely rapidly; however, the resulting system is extremely sensitive to high-frequency disturbances. On the other hand, if the eigenvalues of the observer are small, the estimate of the state approaches the actual value of the state very slowly and the overall performance of the system is substantially degraded. In practice the eigenvalues of the observer are selected to be slightly larger than the largest eigenvalues of the original system. From (21), the eigenvalues of the observer are λ_1 and λ_2. There appears to be no reason for having λ_1 differ from λ_2; therefore, λ_1 and λ_2 were set equal to one another and were made slightly larger than the largest eigenvalue of the original system. The values of the constants β, g_{12}, and g_{22} appear have no appreciable affect on the dynamic response of the vehicle and hence were set equal to one.

4.2 Effects of the Observer on Dynamic Response

The response of the vehicle with the observer was determined as a function of q_1/q_3 for zero initial headway error and a headwind three times the vehicle nominal velocity and for an initial headway error of ten percent and zero headwind. The propulsion system time constant was assumed to be one tenth the headway time; however, Garrard and Kornhauser [2] have demonstrated

that the dynamic response of the optimally controlled vehicle is relatively insensitive to this parameter. The parameter $\frac{2C_D H}{M}$ was set at .025, a value typical of small automated transit vehicles [6,7]. The values of the non-dimensional gains are given in Table I.

The headway error, acceleration, jerk, maximum power, and the time to reduce the headway error to ten percent of its initial value were plotted versus the ratio of q_1 to q_3 for the case in which all states were sensed and for the case in which only position and velocity were sensed. The maximum acceleration and jerk were determined for an initial headway error of ten percent and zero headwind. The maximum headway error and maximum power were determined for zero initial headway and a headwind three times the nominal velocity of the vehicle. The responses obtained for an initial headway error were graphically indistinguishable [Figs. 3-5] and only small increases in maximum headway error and peak power resulted from use of the observer to estimate the acceleration and rate of change of propulsive force.

5. EXAMPLE: DESIGN OF A LONGITUDINAL FEEDBACK CONTROL SYSTEM FOR A HIGH-CAPACITY PRT SYSTEM

A longitudinal control system was designed for a PRT system with the following specifications:

Nominal Mainline Velocity = 50 ft/sec

Minimum Headway Time = 0.5 sec

Vehicle and Passenger Weight = 3200 lbs

Vehicle Length = 10 ft

Maximum Acceleration in Mainline Operation = 4 ft/sec^2

Maximum Acceleration for Merging and Manuevering = 8 ft/sec^2

Maximum Emergency Deceleration = 25 ft/sec^2

Maximum Jerk in Mainline Operation = 4 ft/sec^3

Maximum Jerk in Merging and Manuevering = 8 ft/sec^3

Maximum Headway Error = 7.5 ft

Propulsion System Time Constant = 0.05 sec

The above specifications are typical of many proposed high-capacity PRT systems and result in a system with a mainline capacity of 7200 vehicles/hr. The minimum nominal spearation between adjacent vehicles is 15 feet, thus even in the case in which the leading vehicle encounters a sudden gust of 150 ft/sec, the minimum separation between the leading and following vehicle is 7.5 feet.

The nominal headway for the system is 25 ft., thus a maximum headway error of thirty percent is allowed. From Fig. 2 it can be seen that this criterion will be satisfied for q_1/q_3 greater than one. The maximum allowable jerk in mainline operation is 4 ft/sec^3 and from Fig. 4*, q_1/q_3 must be less than or equal to one. Thus passenger, comfort and headway control requirements are satisfied for q_1/q_3 equal to one. The non-dimensional feedback gains for this case are given in Table 1. The most negative eigenvalue for this system is -10.48; therefore, the eigenvalues of the observers were set at -12.

---

* Non-dimensional jerk is obtained by multiplying dimensional jerk by T^2/V and non-dimensional acceleration is obtained by multiplying dimensional acceleration by T/V.

q_1/q_3	K_{14}	K_{24}	K_{34}	K_{44}
10^{-2}	0.316	2.469	7.776	1.205
10^{-1}	1.000	5.823	11.665	1.547
1	3.162	14.866	18.824	2.151
10	10.000	40.664	32.340	3.217
10^2	31.624	117.204	58.687	5.079
10^3	100.000	349.970	111.879	8.270
10^4	316.243	1068.284	222.545	13.563
10^5	1000.000	3304.094	457.215	22.012
10^6	3162.434	10299.756	959.351	35.050

Table I - Non-Dimensional Gains for Various Values of q_1/q_3, $\tau = .1T$, $q_2 = 10q_1$, $q_3 = 10$, and $q_4 = 10$

The nonlinear vehicle dynamics (6) the dynamics of the propulsion system (7), and the observer (18) were simulated on a digital computer. The following situations are illustrated:

1. Mainline operation with a suddenly applied headwind of 50 ft/sec (Figs 7 and 8)
2. Mainline operation with an initial headway error of 2.5 ft (Fig. 9)
3. An emergency stop with a constant deceleration of 25 ft/sec (Figs 10 and 11)
4. Merging from an off-line station following a trapezoidal acceleration profile (Fig 12)

In Fig. 7 it can be seen that the maximum headway error due to a suddenly applied wind gust of 50 ft/sec is 1.5 ft and the maximum velocity error is 1.8 ft/sec. The response of the vehicle with only position and velocity measured is very near the response with all states sensed. The rapidity with which the estimated acceleration approaches the actual acceleration is illustrated in Fig. 8. As shown in Fig. 9, the maximum jerk for an initial error of 2.5 ft is 4 ft/sec^3, and the maximum acceleration is 0.75 ft/sec^2, thus the resulting ride is not uncomfortable. In this case the difference between the response of the vehicle with all states sensed and the response with only position and velocity sensed are indistinguishable.

During merging, emergency stopping, and maneuvering the vehicle is not operating at mainline conditions. For example, during merging the vehicle starts with zero velocity and accelerates

to line velocity following a commanded acceleration profile as shown in Fig. 12. The differential equations describing the state of the system during merging, emergency stopping, and manuevering are nonlinear with time-varying coefficients. The control system in this study was designed on the basis of state equations which were linear with constant coefficients. However as can be seen from an examination of Figs. 10-12, the resulting controller follows the desired profiles very well even though the conditions are vastly different from those encountered during mainline operation. There is little difference in the dynamic response for the system with the observer and the system in which all states are sensed. Thus a controller designed for mainline operation can also be used for other operations such as emergency stopping, merging, and manuevering. This is important since it shows that a linear controller with fixed gains is sufficient for all control operations and thus the complexity of the control system is minimal.

A conventional filter was also tested as a means of implementing the optimal control without sensing the acceleration and rate of change of propulsive force. The non-dimensionalized velocity measurement was fed into a filter with the following transfer function

$$\frac{\dot{y}}{\theta} = \frac{(\tau_1 s + 1)(\tau_3 s + 1)}{(\tau_2 s + 1)(\tau_4 s + 1)} \qquad (25)$$

where $\tau_1\tau_3 = K_{44}$, $\tau_1 + \tau_3 = K_{34}$, $\tau_2 = .1\tau_1$, and $\tau_4 = .1\tau_3$.*

---

* The notation "s" denotes the Laplace operator.

The output of the filter was used as part of the control input. The resulting approximation of the optimal control was

$$u = -K_{14}y - (K_{24} - 1)\dot{y} - \theta \qquad (26)$$

In this example, use of this filter yielded a controller which approximated the optimal control fairly closely. A typical result is shown in Fig. 9. However, for values of the ratio q_1/q_3 greater than ten, stability problems resulted from use of the filter described above.

The velocity of the vehicle can easily be measured continuously by use of a tachometer; however, it appears that position can only be determined at discrete intervals by means of sensors imbedded in the guideway [6]. The effects of sampling of position on the dynamic response of the system was determined for sensors placed at five-foot intervals on the guideway. At a nominal line velocity of 50 feet/sec, this corresponds to a sampling time of one-tenth of a second. Of course when the vehicle is traveling at less than line velocity, for example during merging or emergency stopping, the sampling rate decreases.

A typical example of vehicle response with sensors placed at five-foot intervals is shown in Figs. 13-16. It can be seen that sampling does not significantly degrade performance. Emergency stopping and response to suddenly applied headwinds were also considered and the vehicle response did not differ significantly from that obtained with continuous position measurement. The results for merging and emergency stopping are summarized in Table II.

Merging						Emergency Stop
Final Headway Error			Final Velocity Error		Maximum Jerk	Final Headway Error
	per cent	ft.	per cent	ft/sec	ft/sec^3	ft.
All States Sensed	1.70	0.425	2.69	1.35	10.40	4.25
Positon and Velocity Only Sensed	1.91	0.475	2.95	1.45	10.66	4.70
Position and Velocity Only Sensed-Position Sampled	2.00	0.500	3.00	1.50	10.48	4.80

Table II Comparison of System Errors for All States Sensed, Position and Velocity Only Sensed, and Position and Velocity Only Sensed-Position Sampled.

Conculsion

From the results presented above, it appears that optimal control theory can be usefully applied to the design of longitudinal control systems for automated transit vehicles with a wide variety of characteristics. The resulting control systems keep headway and velocity errors small without causing passenger discomfort, and excellent dynamic response is achieved during mainline operation, merging and demerging, maneuvering and

emergency stopping. The controllers are linear with constant gains and should be relatively economical to implement and simple enough to insure reliability. It is shown that little degradation in dynamic response results if only position and velocity errors are measured and system performance is little effected by reasonably frequent sampling of position error. The effects of propulsion system nonlinearties, uncertain knowledge of vehicle parameters, random sensor errors, and random disturbances are currently being investigated.

Also the dynamics of slippage between the wheels and guideway has not been included in the current simulation and the effects of sensor spacings and sensor errors has not been fully investigated. These effects can be significant during emergency stopping situations when high decelerations are commended and rubber-tired vehicles are used.

ACKNOWLEDGEMENT

This research was partially supported by grant Minn UTR-3(71) of the Urban Mass Transit Administration of the United States Department of Transportation. The authors also gratefully acknowledge the assistance of the Systems and Research Division of Honeywell, Inc., St. Paul, Minnesota in providing the computer program used for solving the matrix Ricatti equation.

REFERENCES

1. Hajdu, L.P., Gardiner, K.W., Tomura, H. and Pressman, G.L., "Design and Control Considerations For Automated Ground Transportation Systems," Proc. IEEE, Vol. 56, No. 1, April 1968, pp. 493-513.

2. Garrard, W.L. and Kornhauser, A.L., "Design of Optimal Feedback Systems for Longitudinal Control of Automated Transit Vehicles," Transportation Research (to appear), 1973.

3. Luenberger, D.G., "An Introduction to Observers," IEEE Trans. Aut. Control, Vol. AC-16, No. 6, Dec. 1971, pp. 596-603.

4. Boyd, R.K. and Lukas, M.P., "How to Run an Automated Transportation System," IEEE Trans. Sys., Man, and Cybernetics, Vol. SMC-2, No. 3, July 1972, pp. 331-341.

5. Munson, A.V., "Quasi-Synchronous Control of High-Capacity PRT Networks," in Personal Rapid Transit, ed. J.E. Anderson, J.L. Dais, W.L. Garrard, and A.L. Kornhauser, Univeristy of Minnesota, 1972, pp. 325-350.

6. Whitney, D.E., and Tomizuka, M., "Normal and Emergency Control of a String of Vehicles by Fixed Reference Sampled-Data Control," in Personal Rapid Transit, ed. J.E. Anderson, J.L. Dais, W.L. Garrard, and A.L. Kornhauser, University of Minnesota. 1972, pp. 383-404.

7. Wilkie, D.F., "A Moving Cell Control Scheme for Automated Transportation Systems," Transpn. Sci, Vol. 4, No. 4, Nov. 1970, pp. 347-364.

8. Lee, E.B. and Markus, L., Foundations of Optimal Control Theory, John Wiley and Son, New York, 1967.

9. Kleinman, D.L., "On an Iterative Technique for Ricatti Equation Computations," IEEE Trans. Aut. Control, Vol. AC-13, No. 1, Feb. 1968, pp. 114-115.

Fig. 1 Block Diagram of Optimal System with Observer

Fig 1 Guidance & Navigation

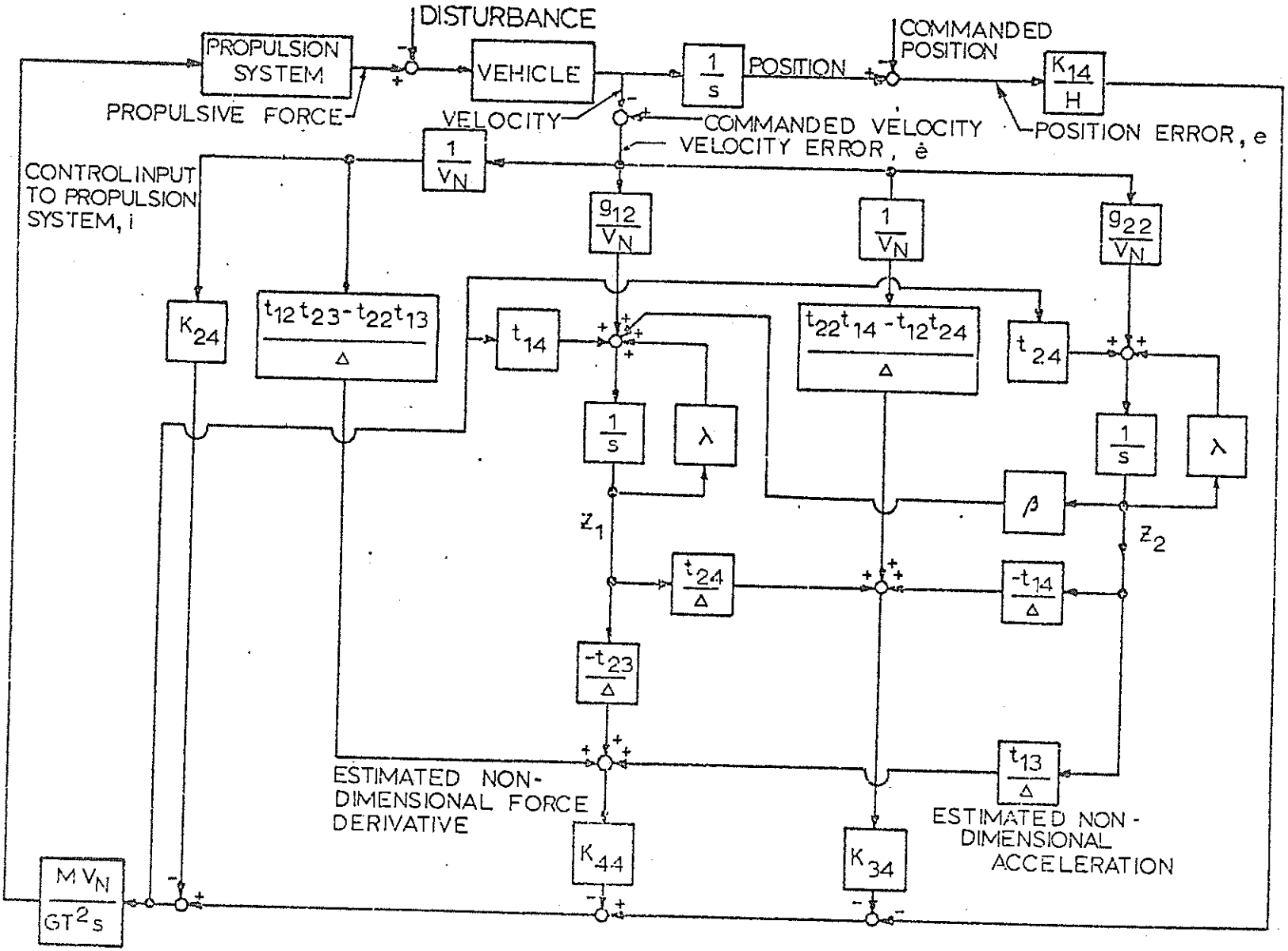


Fig. 2 Maximum Percentage Headway Error vs. q_1/q_3 for a Headwind Three Times Nominal Velocity: $q_2 = 10q_1$, $q_4 = 10$ $q_3 = 10$ and $\tau = .1T$

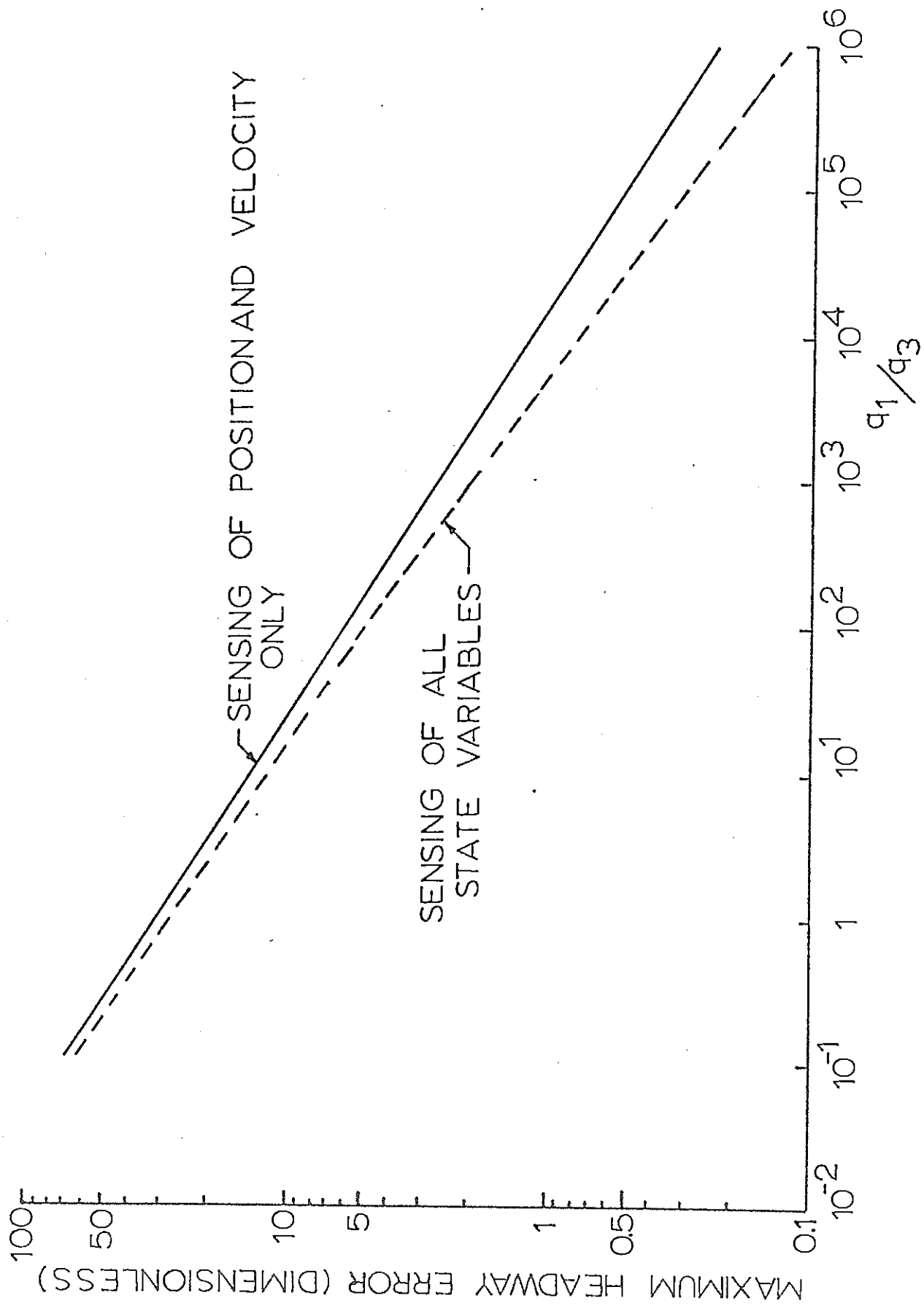

MAXIMUM HEADWAY ERROR (DIMENSIONLESS)
100
50
10
5
1
0.5
0.1
10^{-2}
10^{-1}
1
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}
10^{6}
q_1/q_3
SENSING OF POSITION AND VELOCITY
SENSING ONLY
SENSING OF ALL STATE VARIABLES

Fig. 3 Maximum Non-dimensional Acceleration Error vs. q_1/q_3 for a Ten Percent Initial Headway Error: $q_2 = 10q_1$, $q_4 = 10$ q_3 = and 10 and $\tau = .1T$

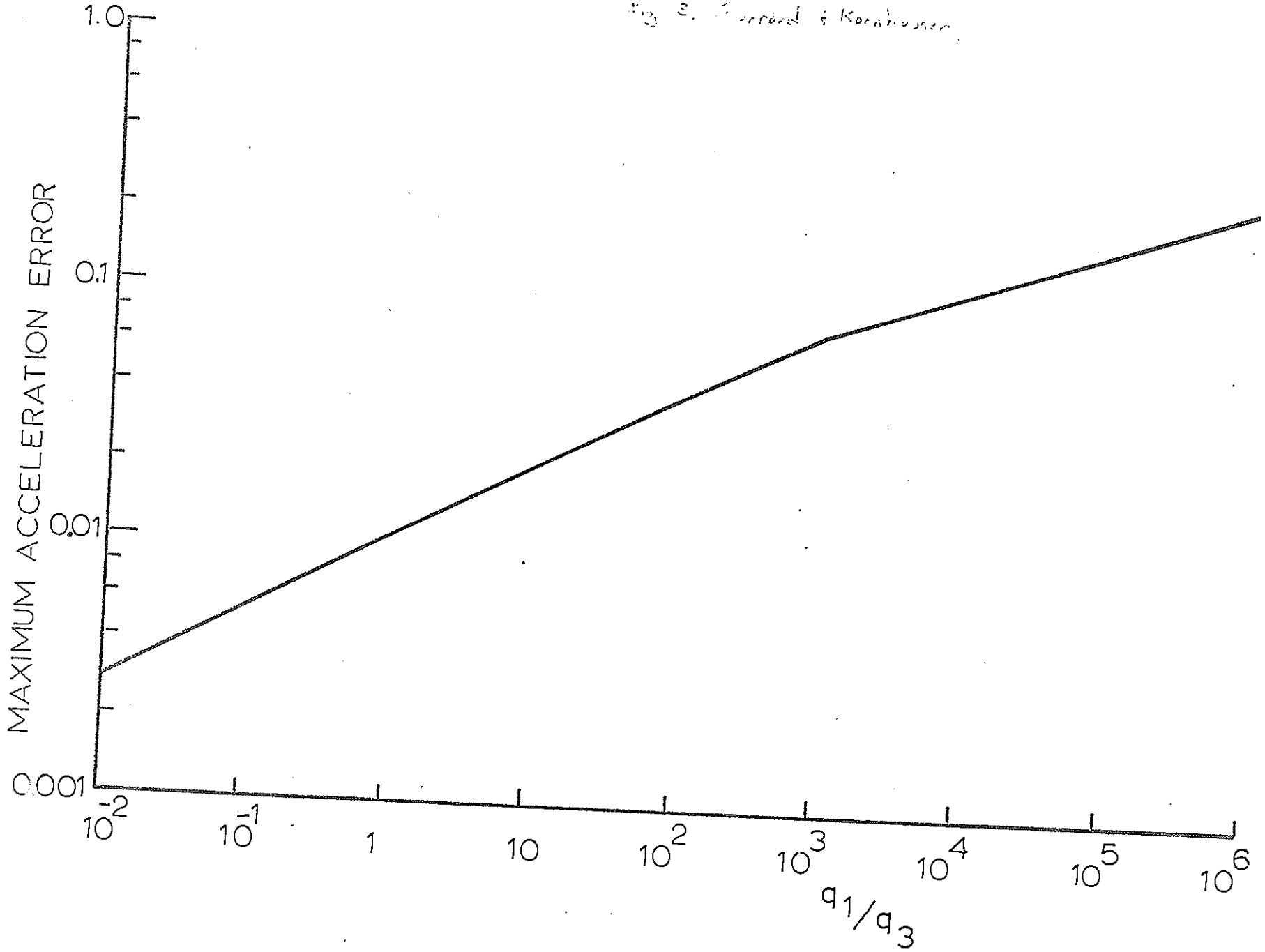

MAXIMUM ACCELERATION ERROR
1.0
0.1
0.01
0.001
10^{-2}
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6
q_1/q_3

Fig. 4 Maximum Non-dimensional Jerk vs. q_1/q_3 for a Ten Percent Initial Headway Error: $q_2 = 10q_1$, $q_4 = 10$ $q_3 = 10$ and $\tau = .1T$

Fig 1 Garrard & Kornhauser

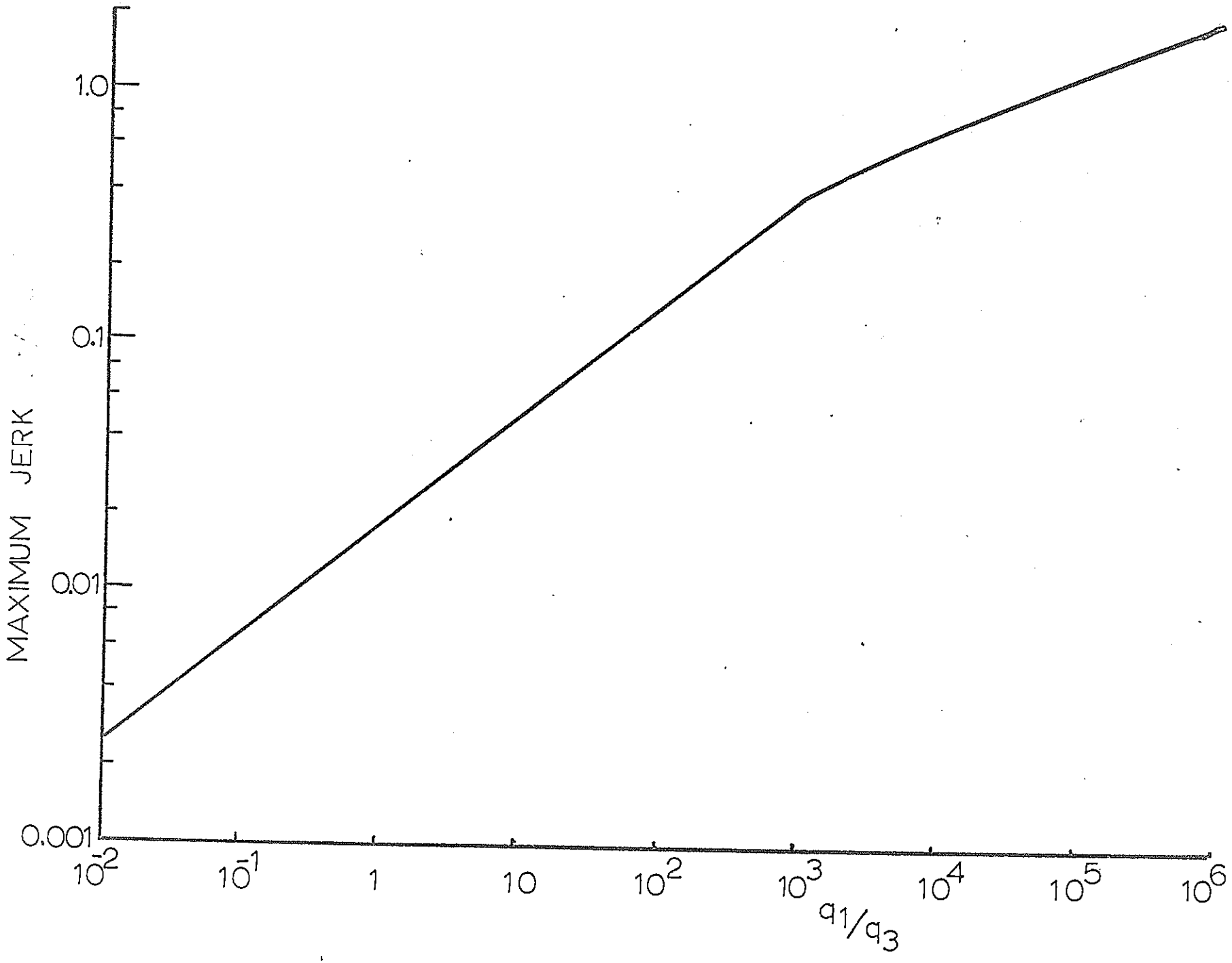


Fig. 5 Non-dimensional Time Required to Reduce Headway Error to Ten Percent of Initial Value vs. q_1/q_3: $q_4 = 10$ $q_2 = 10q_1$, $q_3 = 10$ and $\tau = .1T$

Fig 5 : Gunnard and Kornhauser

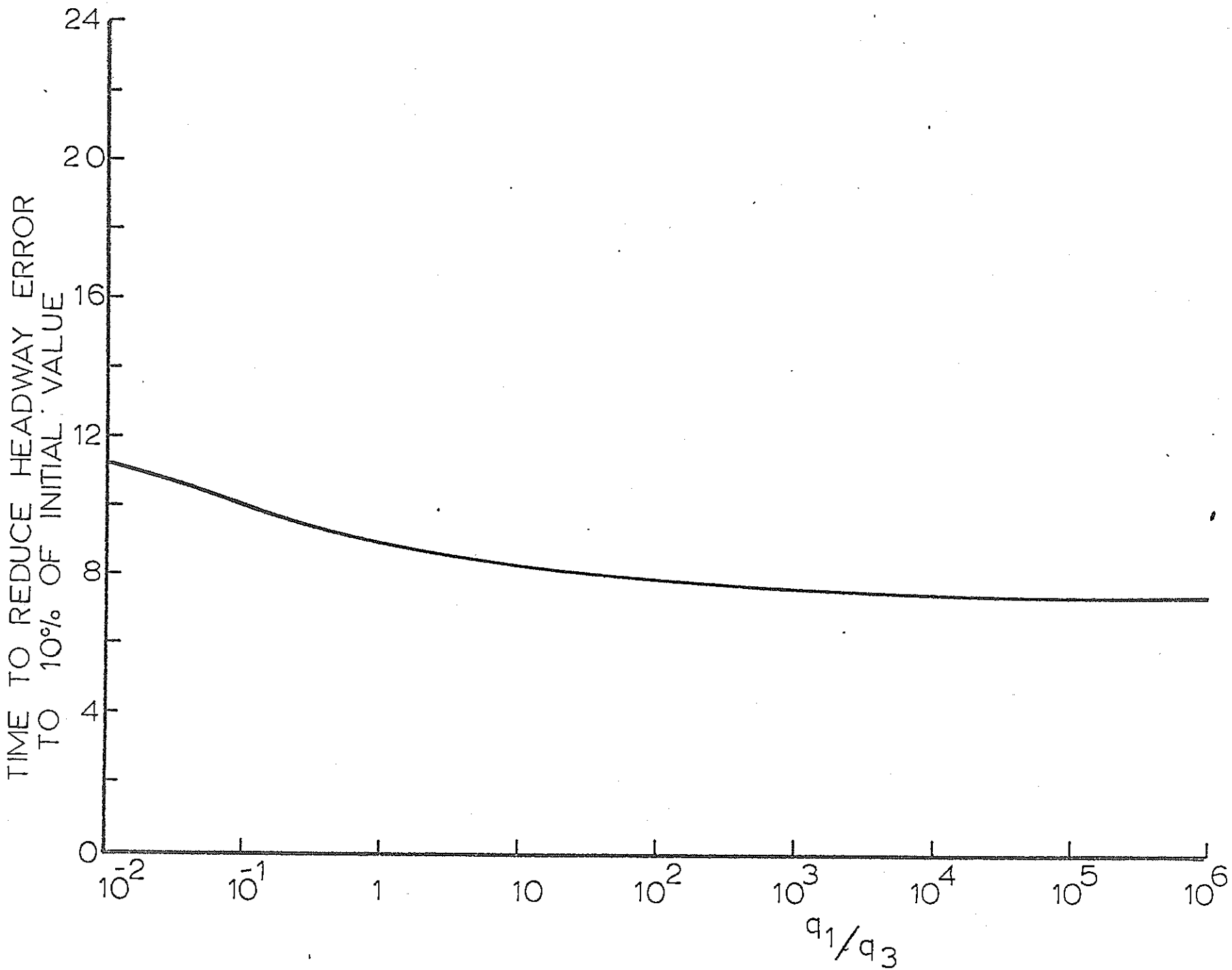


Fig. 6 Peak Dimensionless Power vs. q_1/q_3 for a Headwind Three Times Nominal Mainline Velocity, $q_4 = 10$, $q_3 = 10$, $q_2 = 10q_1$ and $\tau = .1T$

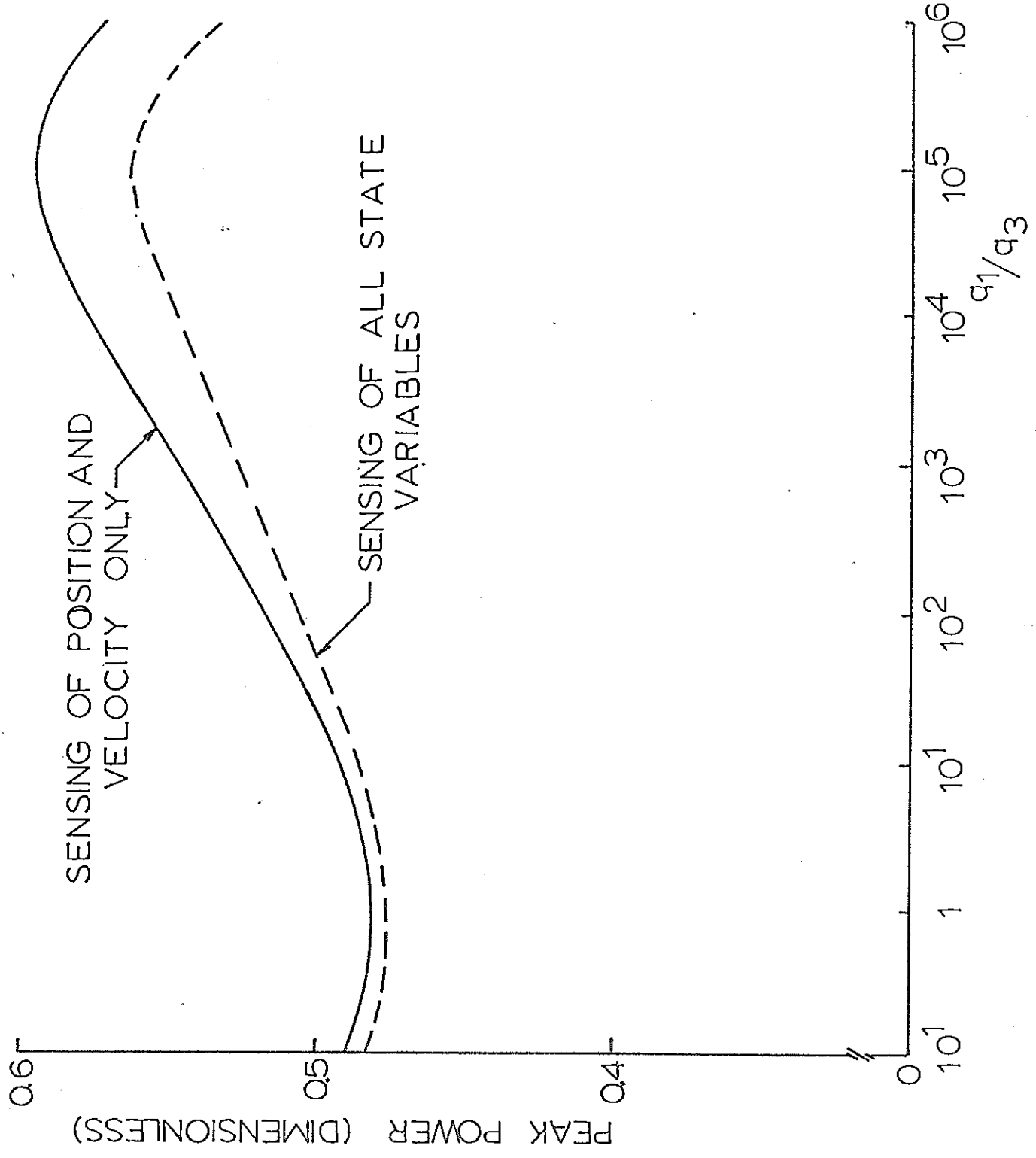


Fig 6 Gjerrard & Kornhauser

Fig. 7 Position and Velocity Error for a 50 ft/sec Headwind-Mainline Operation

Fig 7 Conrad & Kornhauser

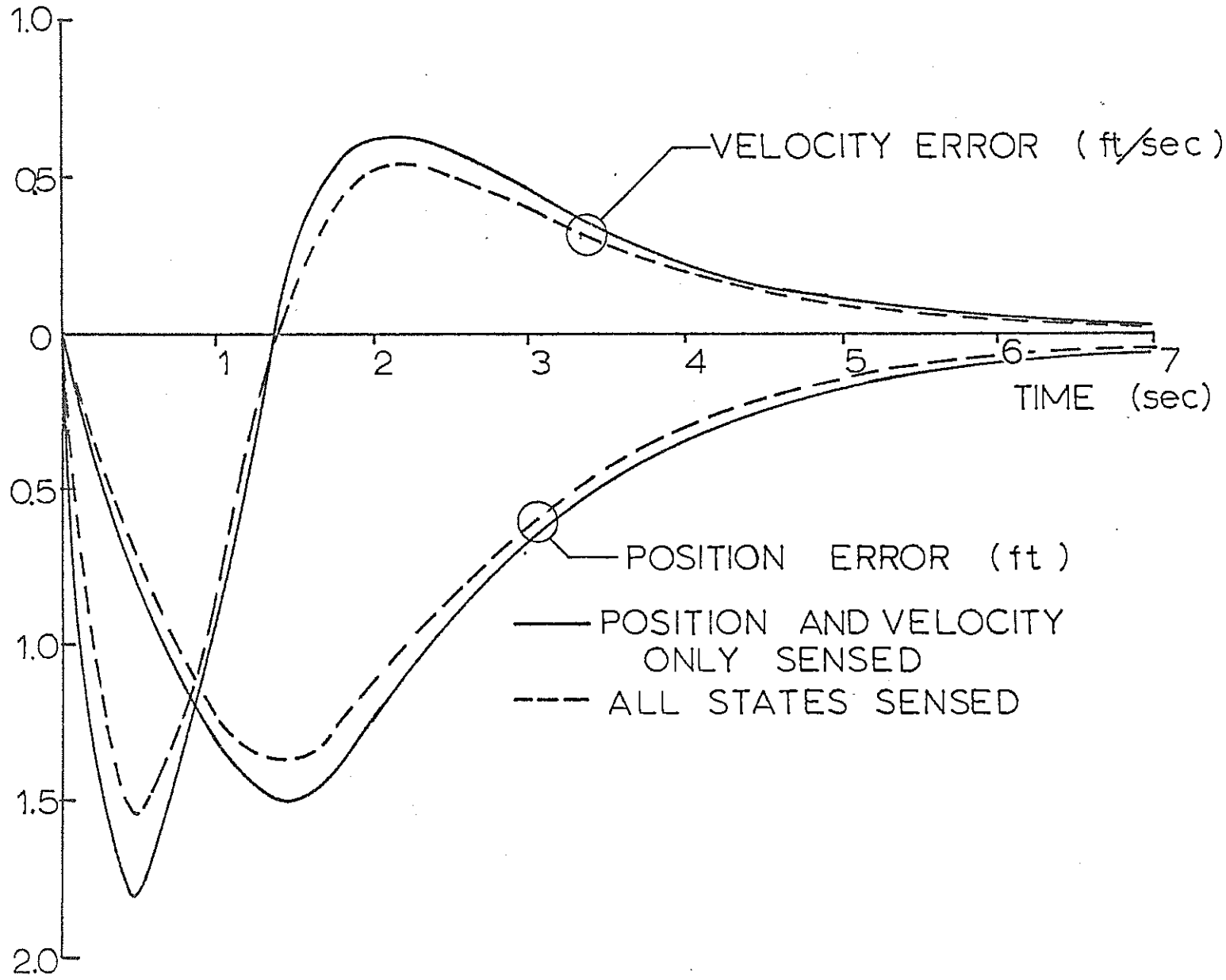


Fig. 8 Actual and Estimated Acceleration for a 50 ft/sec Headwind-Mainline Operation

Fig 8 Garrard and Kornhauser

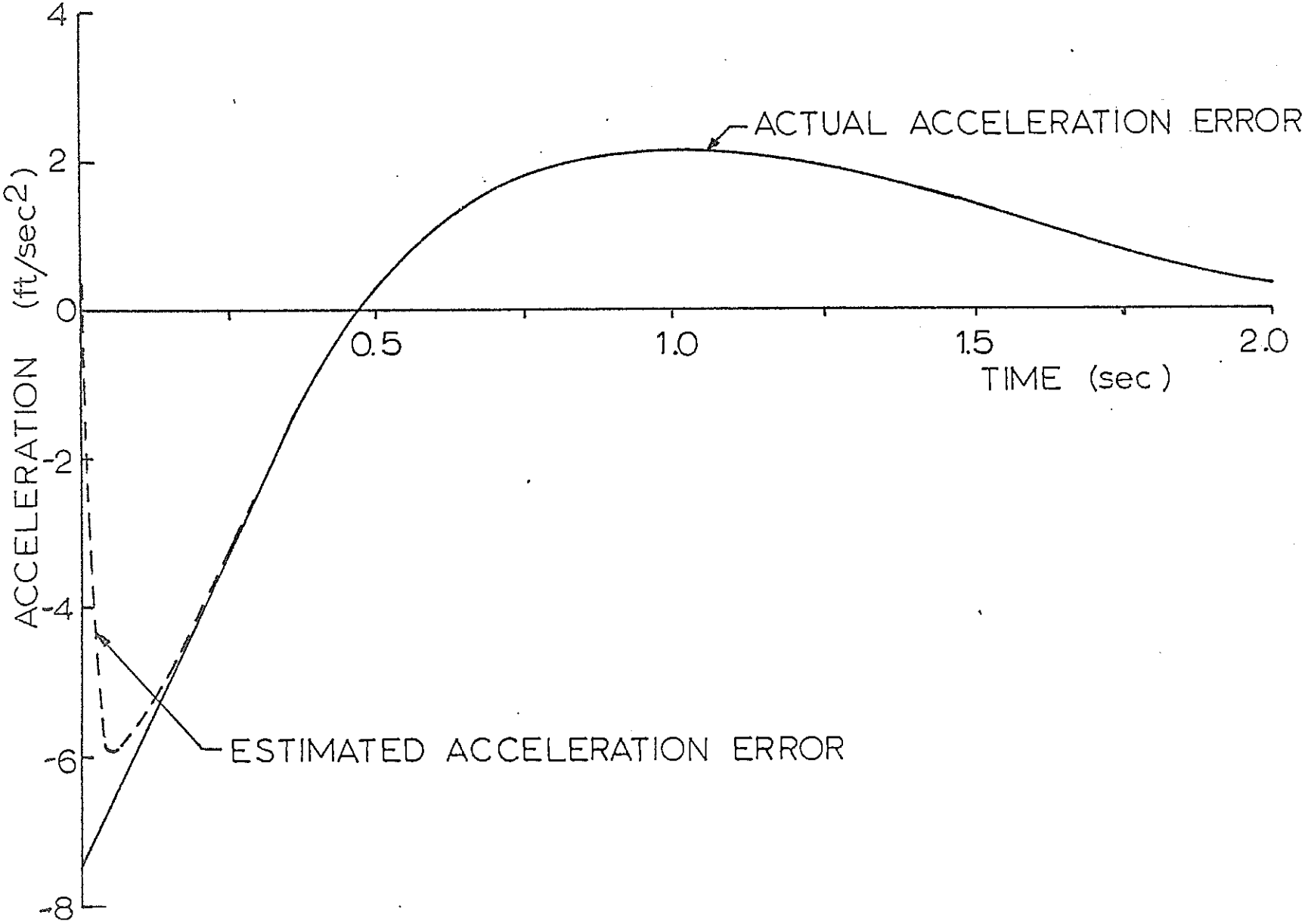


Fig. 9 Position Error, Velocity Error, Acceleration, and Jerk for a 2.5 ft Initial Headway Error-Mainline Operation

Fig 9 Gerrard & Kornhauser

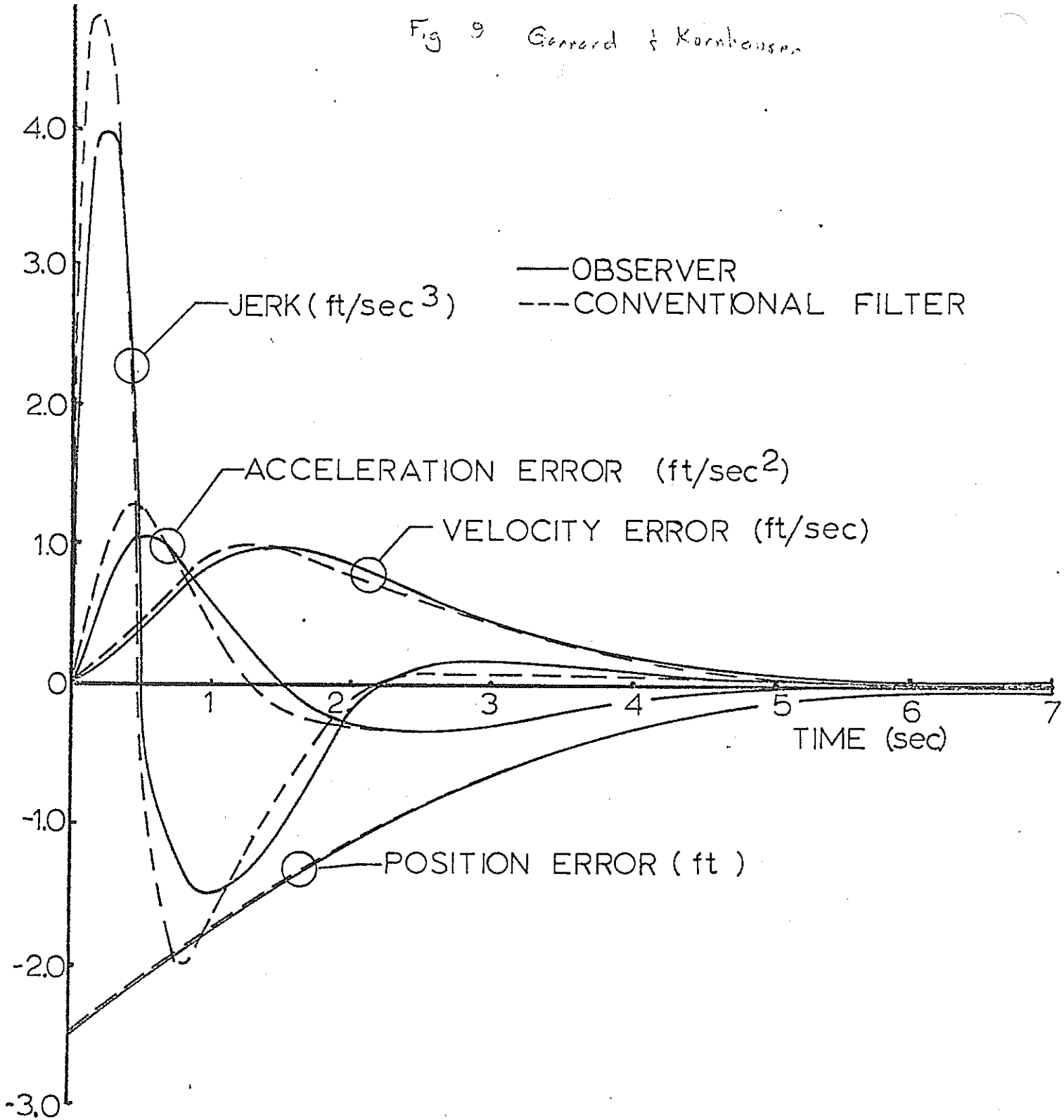


Fig. 10 Velocity - Emergency Stopping

Fig 10 Garrard & Kornhauser

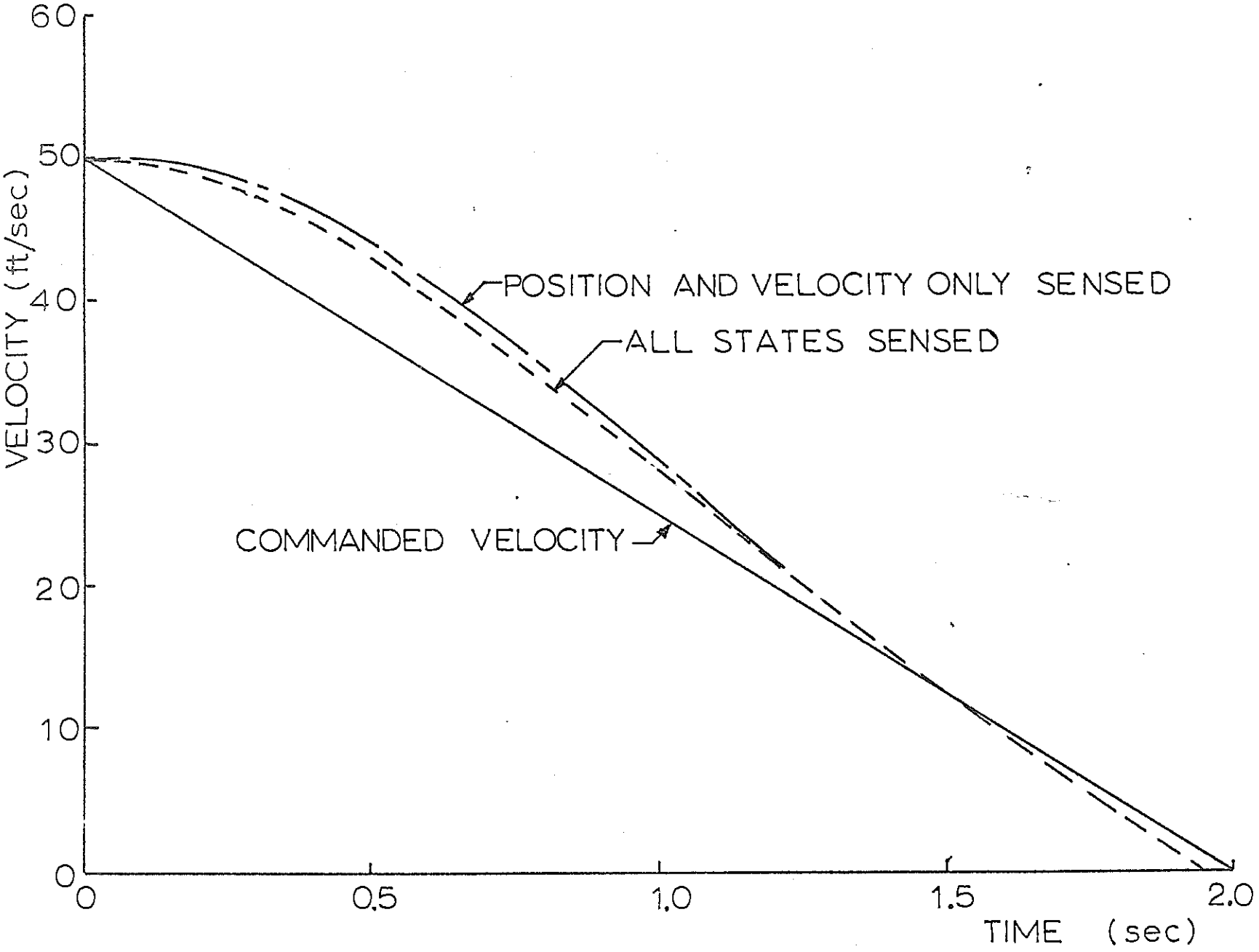


Fig. 11 Position - Emergency Stopping

Fig 11 Command Response

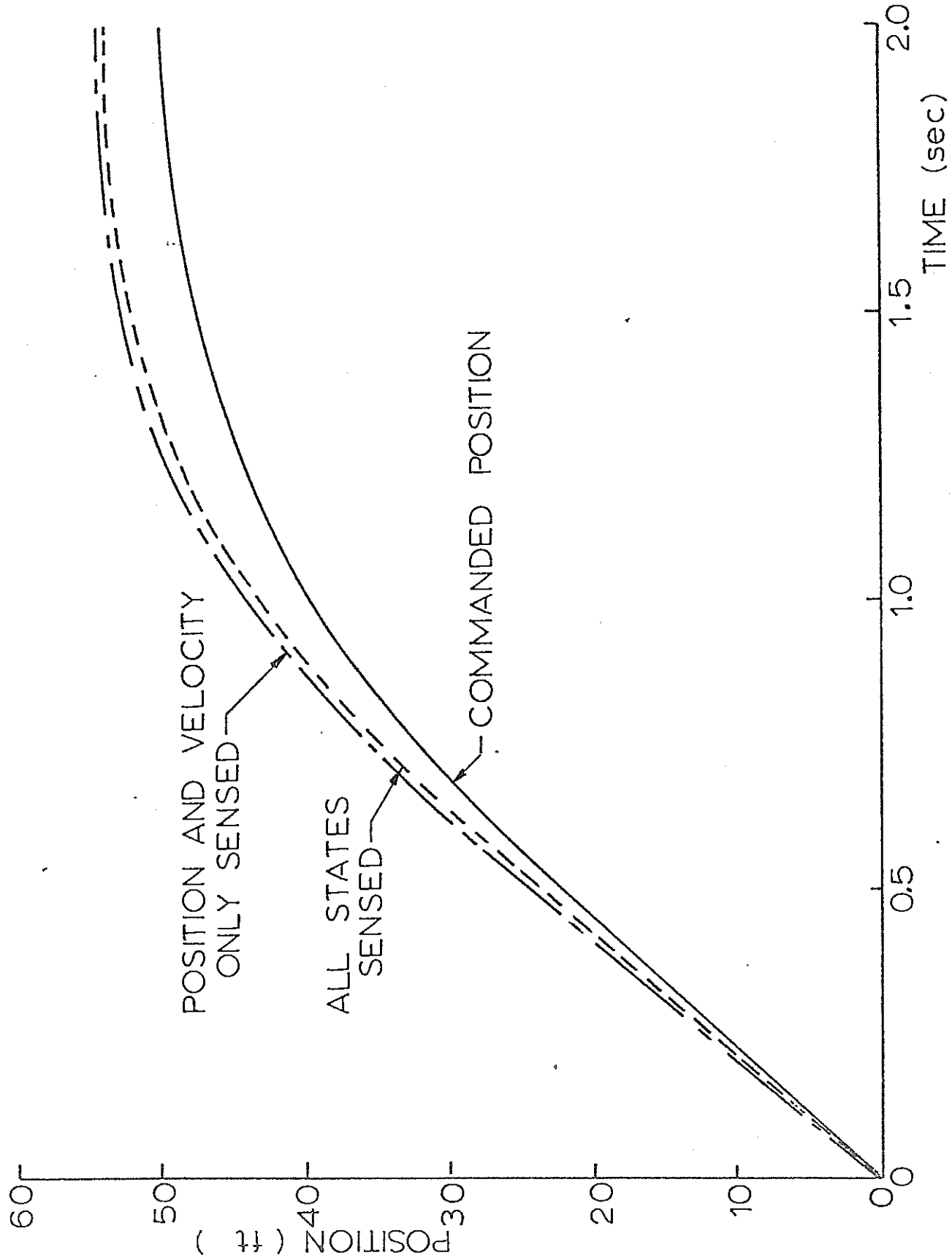


Fig. 12 Acceleration - Merging from Off-Line Station

Fig 12 Gennard & Kornhauser

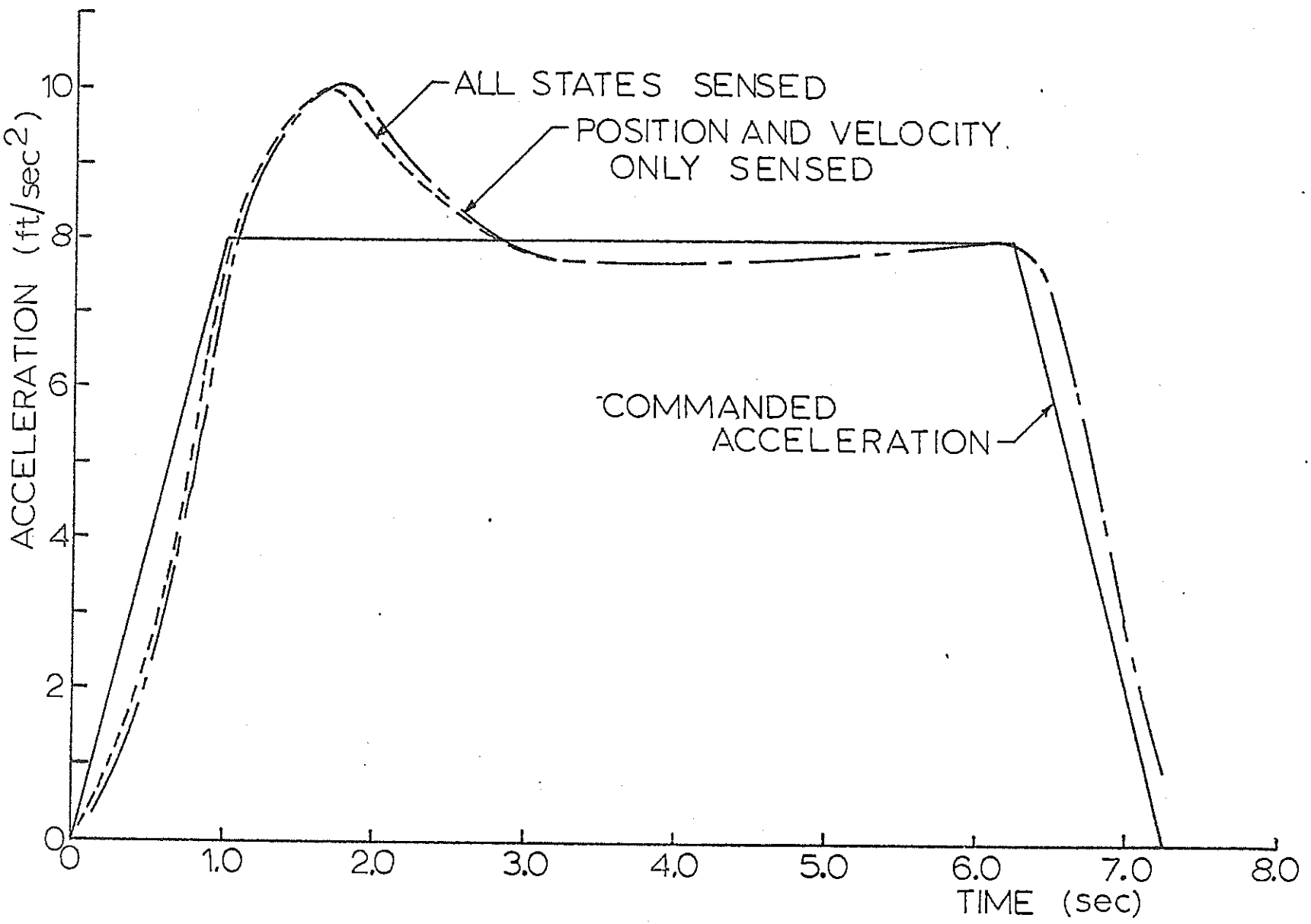


Fig. 13 Position and Velocity Errors for a 2.5 Foot Initial Position Error - Sampled and Continuous Data

Fig 13 Gerrard & Kornhauser

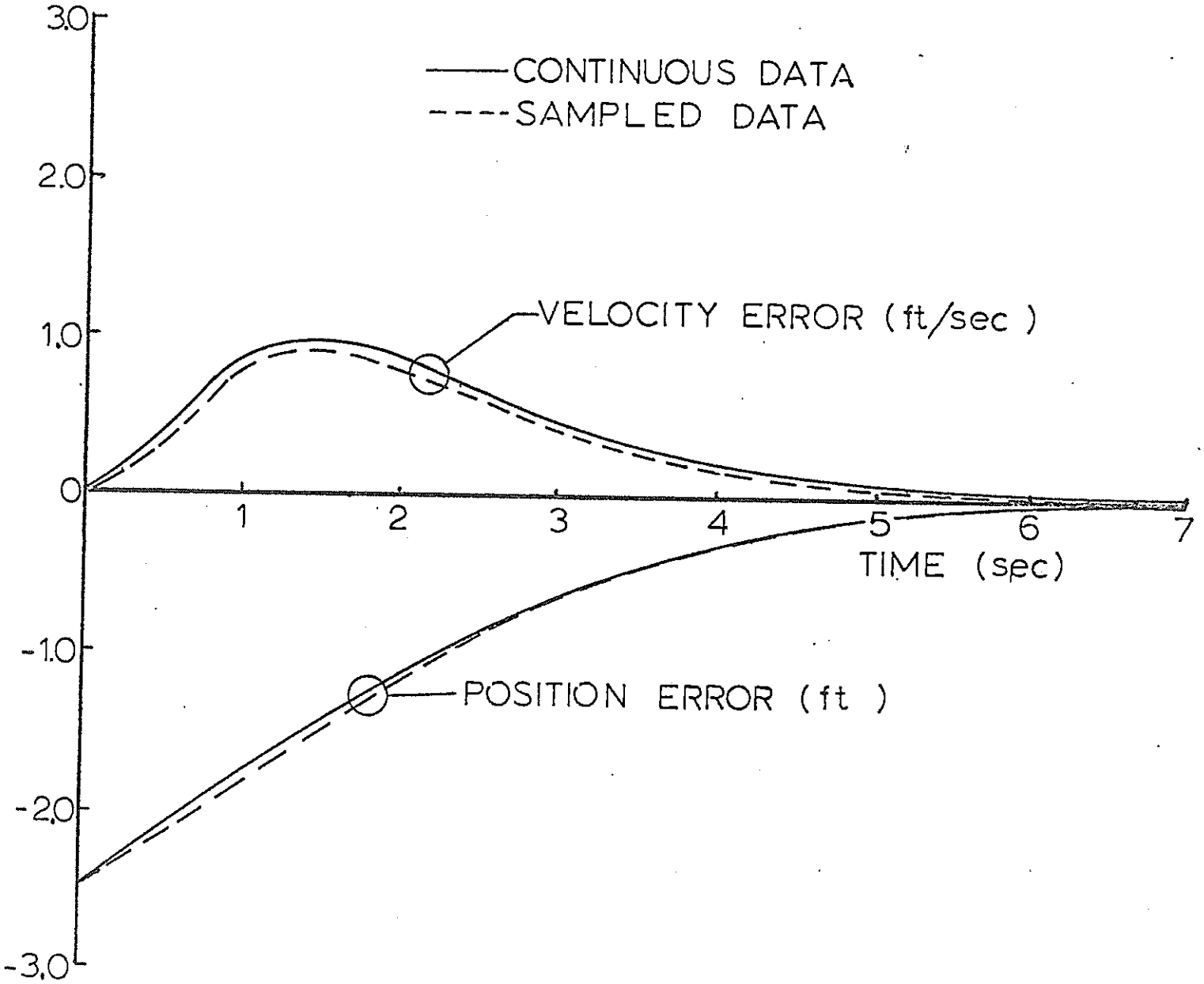


Fig. 14 Acceleration for a 2.5 Foot Initial Position Error - Sampled and Continuous Data

Fig 14 Stanard & Karnahan

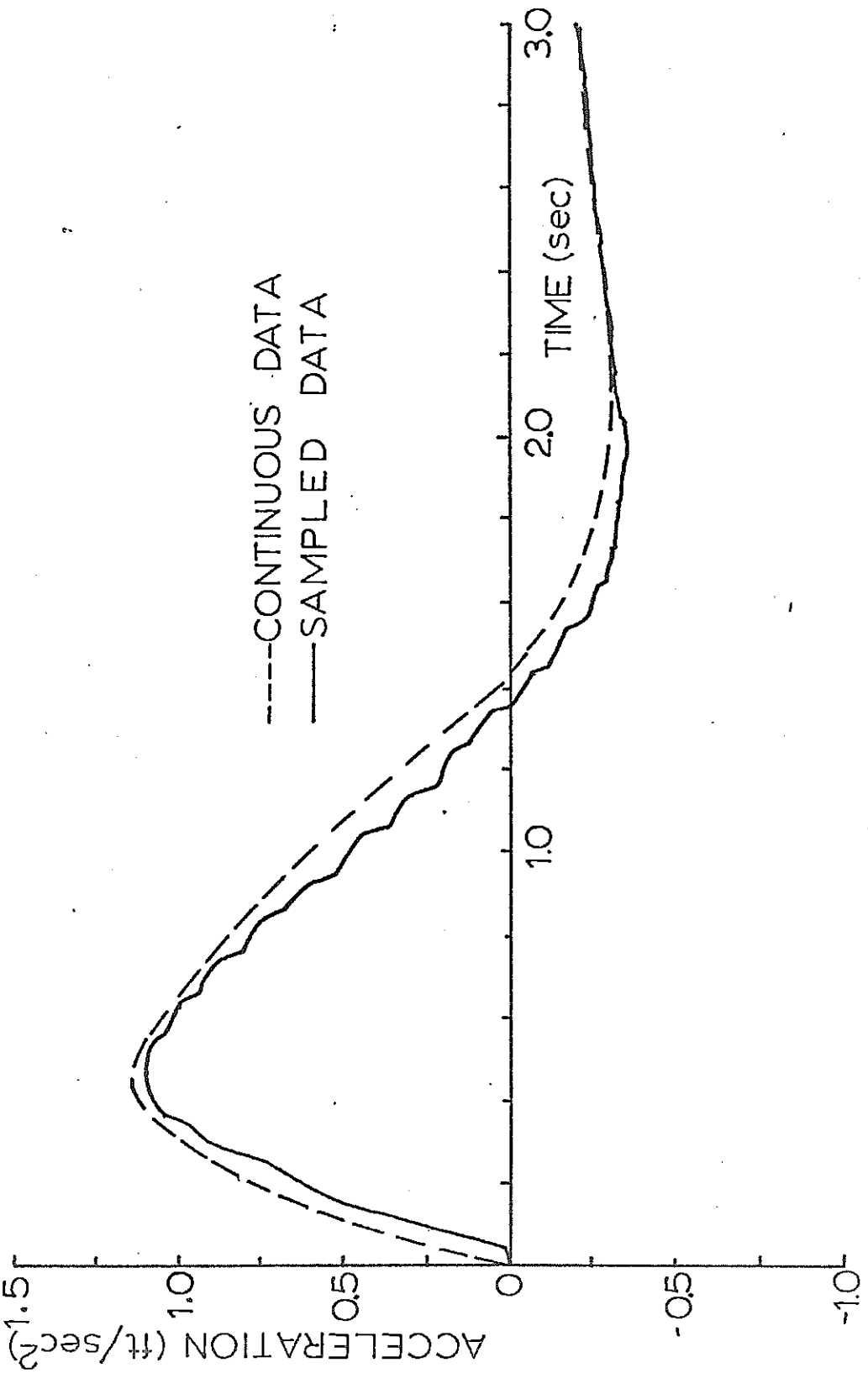


Fig. 15 Jerk for a 2.5 Foot Initial Position Error - Sampled and Continuous Data

Fig 15 Garrard & Kornhauser

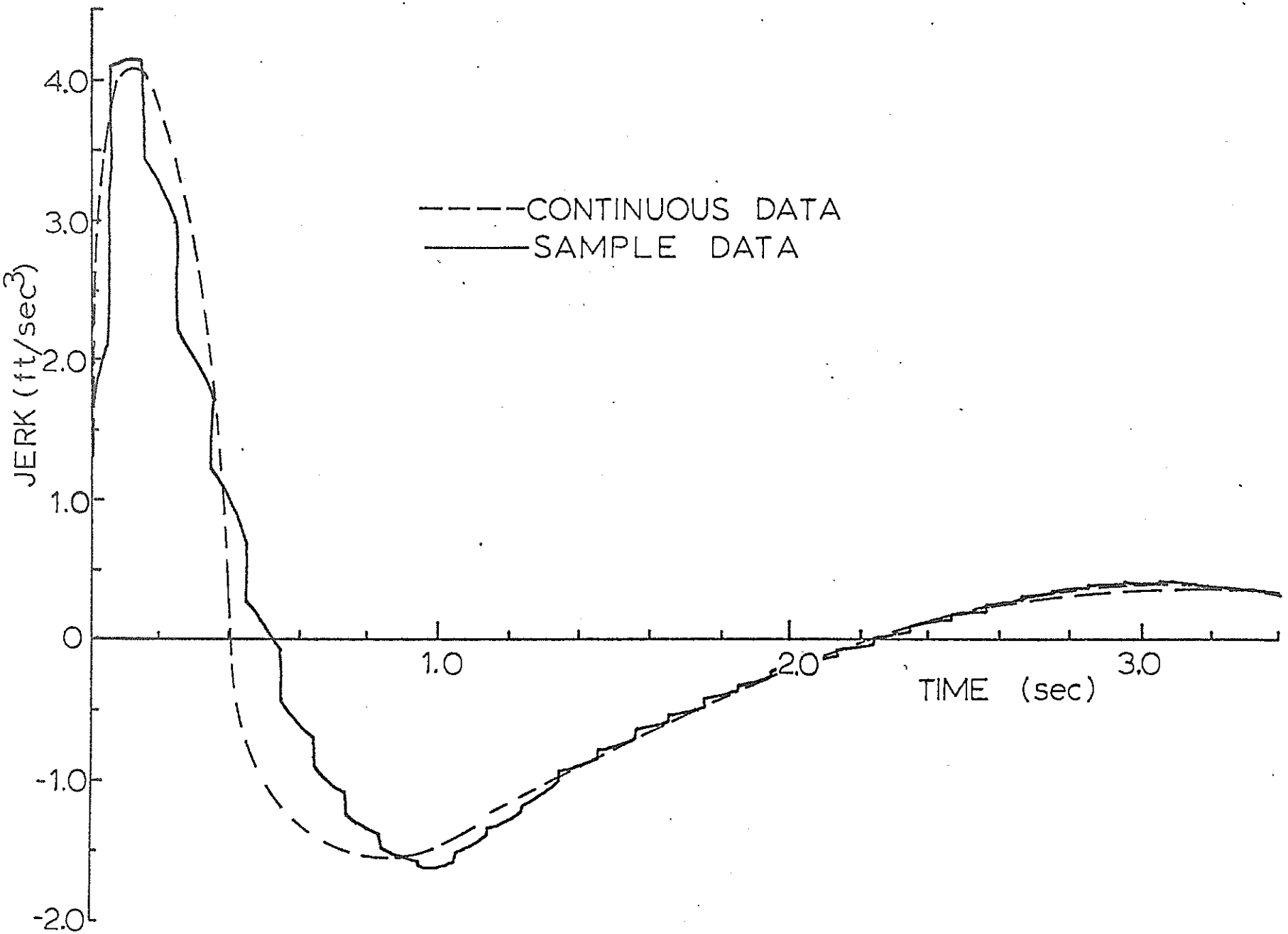


Fig. 16 Acceleration - Merging from Off-Line Station - Sampled and Continuous Data

Fig 16 Garrard & Kornhauser