control inputs at each step. Therefore, for every such state, $\mathbf{x}(k)$, the control input of polarity $\mu(k) = -\zeta_1$, and duration $\tau_k = \delta_1$, is optimal if and only if, for every N, the state $\mathbf{x}(k+1)$ at the next step is in Δ_{n-1}. Using (18) with $\mu(k) = -\zeta_1$, $\tau_k = \delta_1$, and $\mathbf{x}(k)$ in the form given in (32), the state at the next step can be written

$$\mathbf{x}(k+1) = \zeta_1 \mathbf{G}(T)\mathbf{w}_1(\delta_1) + \mathbf{G}(T)\sum_{i=2}^{N} \zeta_i \mathbf{w}_i(\delta_i) + \zeta_1 \mathbf{G}(T)\mathbf{h}(-\delta_1) \quad (33)$$

Using (5), (6), (8), and (19).

$$\mathbf{G}(T)\mathbf{w}_1(\delta_1) = -\mathbf{G}(T)\mathbf{h}(-\delta_1)$$

Therefore (33) reduces to

$$\mathbf{x}(k+1) = \mathbf{G}(T)\sum_{i=2}^{N} \zeta_i \mathbf{w}_i(\delta_i)$$

From (5) and (19) it follows that

$$\mathbf{G}(T)\mathbf{w}_i(\delta_1) = \mathbf{w}_{i-1}(\delta_{i-1})$$

Therefore the state at the $(k+1)$ sampling instant is given by

$$\mathbf{x}(k+1) = \sum_{i=1}^{N-1} \zeta_i \mathbf{w}_i(\delta_i)$$

which, from Lemma 2, is in Δ_{N-1}. If $\mathbf{x}(k)$ is in Δ_N but not in Δ_{N-1}, then the foregoing input insures that $\mathbf{x}(k+1)$ is in Δ_{N-1}. It cannot also be in Δ_{N-2}, since the optimal control sequence can take $\mathbf{x}(k)$ into the origin in exactly N steps and no less. Therefore if $\mathbf{x}(k)$ is in Δ_N but not in Δ_{N-1} the control input specified by Theorem 2 insures that $\mathbf{x}(k+1)$ is in Δ_{N-1} but not in Δ_{N-2}. By repetition of the foregoing computation, it follows that, at every step, the control input of polarity $\mu = -\zeta_1$ and duration $\tau = \delta_1$ is optimal.

References

1 W. L. Nelson, "Pulse-Width Relay Control in Sampling Systems," Journal of Basic Engineering, Trans. ASME, Series D, vol. 83, 1961, pp. 65–76.

2 R. E. Kalman and J. E. Bertram, "Control System Analysis and Design Via the 'Second Method' of Lyapunov," Journal of Basic Engineering, Trans. ASME, Series D, vol. 82, 1960, pp. 371–400.

3 R. E. Kalman and J. E. Bertram, "General Synthesis Procedure for Computer Control of Single and Multiloop Linear Systems," *Trans. AIEE*, vol. 77, part II, 1958, pp. 602–609.

4 W. L. Nelson, "Pulse Width Control of Sampled Data Systems," Technical Report T-35/B, Dept. of Elect. Engineering, Columbia University, New York, N. Y., July, 1959.

5 R. E. Kalman, "Optimal Control of Saturating Systems by Intermittent Action," IRE WESCON Conv. Record, vol. 1, Part 4, 1957, pp. 130–135.

6 R. E. Kalman, "On the General Theory of Control Systems," Proc. First International Congress on Automatic Control, Moscow, USSR, 1960.

7 C. A. Desoer and J. Wing, "An Optimal Strategy for a Saturating Sampled-Data System," *I. R. E. Trans. on Automatic Control*, vol. AC-6, February, 1961, pp. 5–15.

8 R. Bellman, I. Glicksberg, and O. Gross, "On the 'Bang-Bang' Control Problem," *Quarterly of Applied Mathematics*, vol. XIV, no. 1, April, 1956.

9 R. E. Kalman "Analysis and Design Principles of Second and Higher Order Saturating Servomechanisms," Appendix II, *Trans. AIEE*, vol. 74, part II, 1955, pp. 294–310.

10 D. W. Bushaw, "Differential Equations With a Discontinuous Forcing Term," PhD thesis, Princeton Univ., Princeton, N. J., 1952.

11 I. Flugge-Lotz, "Discontinuous Automatic Control" (book), Princeton University Press, Princeton, N. J., 1953.

DISCUSSION

Charles A. Desoer[5] and J. Wing[6]

The author should be commended for his interesting paper on the minimal time regulator problem. In particular, Dr. Nelson presents a very interesting and informative example of a second-order underdamped plant whose input is amplitude saturating. He actually establishes an optimal control for such a plant. Now, as is well known in the case of an n-dimensional state space, the optimal control is not unique for sampled data systems except for states in Γ_n and for a set of states of measure zero outside Γ_n. The author is aware of this since he uses the expression "at least" in conclusion (*d*) of Theorem 1. It may be worthwhile to point out that the author's example is a special case of a general theory recently developed by the discussers.[7] In Dr. Nelson's example what is defined in the general theory as the "critical hypersurface" reduces to a straight line; namely, the line through the origin in the direction of $\mathbf{v}_2$ in Fig. 4(*a*).

The discussers would like to clarify a point concerning Theorem 1. In the proof of this theorem Dr. Nelson has to use an assumption that is not stated in the theorem itself. This assumption is to the effect that the forcing vector $\mathbf{b}$ of (1) must have all nonzero components with respect to the basis that diagonalizes $\mathbf{A}$. In other words, the input μ must be coupled to every mode of the system. With this assumption Theorem 1, as far as conclusions (*a*) and (*b*) are concerned, gives a set of *sufficient* conditions for complete controllability with admissible controls. Examination of the author's Appendix (proof of Theorem 1) establishes the following result: when $\mathbf{G}(T)$ has distinct eigenvalues, the above-mentioned condition on $\mathbf{b}$, and the restriction of the eigenvalues of $\mathbf{G}(T)$ to the closed unit circle are both *necessary* and *sufficient* conditions for complete controllability with admissible controls.

What happens when $\mathbf{G}(T)$ has multiple eigenvalues? The only set of necessary and sufficient conditions for complete controllability with admissible controls known to the discussers, and proved in the reference mentioned, is given by the following theorem:

The system (10) *is controllable with admissible controls*, if and only if (i) the vectors $\mathbf{v}_k = \mathbf{G}(T)^{-k}\mathbf{h}(T)$; $(k = 1, 2, \ldots, n)$ are linearly independent; (ii) the eigenvalues of $\mathbf{G}(T)$ are in the closed unit circle (i.e., $|z_k| \leq 1$).

Let us point out a simple fact. Given an arbitrary matrix $\mathbf{G}(T)$ satisfying condition (ii), it may happen that there does not exist a vector $\mathbf{h}(T)$ such that condition (i) can be satisfied. This would be the case if $\mathbf{G}(T)$ were a multiple of the unit matrix, say, $\mathbf{G}(T) = \alpha\mathbf{I}$ where $0 < \alpha \leq 1$. In his conclusion, Dr. Nelson mentions "the conceptual difficulty of determining the optimal control regions in higher dimensional state spaces." The discussers do not believe that this is so difficult, since in their general theory mentioned earlier, an arbitrary nth order system governed by a difference equation of form (10) is considered. The system is arbitrary in the sense that the only constraint on (10) is that it be controllable by admissible controls. Under this general condition, the discussers are able to specify a unique critical hypersurface which partitions the state space in two parts and has the important property that an optimal control law is determined by a simple nonlinear function of the distance from the critical hypersurface to the state of the system. That is to say, the optimal control law has been reduced to a conceptually simple functional defined on the state space. For any initial state $\mathbf{x}(0)$,

[5] Associate Professor of Electrical Engineering, University of California, Berkeley, Calif.

[6] Associate in Electrical Engineering, University of California, Berkeley, Calif.

[7] C. A. Desoer and J. Wing, "Minimal Time Regulator Problem for Linear Sampled-Data Systems (General Theory)," ERL Report 346, Feb. 14, 1961, University of California, Berkeley. To appear in the September, 1961, issue of the *Journal of the Franklin Institute*.

this optimal control law will generate one by one some appropriate $\gamma_i(k)$'s which appear in conclusion (*d*) of Theorem 1.

Author's Closure

The author appreciates the valuable discussion of Professor Desoer and Mr. Wing. The discussers are correct in their point concerning Theorem 1. To be rigorous, the statement of Theorem 1 should include the condition that the input, μ, is coupled to every mode of the system or, equivalently, that there is no zero-pole cancellation in the transfer function of the plant. The other sets of conditions given by the discussers for complete controllability with admissible controls are also a useful supplement to the material in the second section of this paper.

After studying the general theory for minimum time control in discrete systems developed by the discussers, it is the author's opinion that the particular example of optimal control for the second-order underdamped plant given in the fourth section of this paper is not a special case of their general theory, but rather is a different but equally valid optimal control strategy. It corresponds to their control law within Γ_2, of course, since in this region (see Fig. 4) the control law is unique. Outside Γ_2 there are many possible optimal strategies. The author selected that one which allowed on-off control to be optimal everywhere outside Γ_2, since optimal on-off control was the main subject of this paper. Since, as the discussers state, the "critical hypersurface" reduces to a straight line through the origin in the direction of $\mathbf{v}_2$, their general theory prescribes the optimal input as a *saturation* function of the distance γ_1, to this line, not as the signum function. The author will resolve this special point through correspondence with the discussers. Their excellent work, given as a reference by the discussers, is recommended to the reader who wishes to pursue this point.

On the general subject of time-optimal control the conceptual difficulty, mentioned in the conclusion of this paper, in specifying a particular optimal strategy in higher dimensional state spaces is now removed by the general theory developed by Desoer and Wing, at least for the case of discrete systems controllable by saturating amplitude controls.

Furthermore, it was shown in this paper that between saturating amplitude control and pulse width on-off control there exists a basic similarity in the topography of the optimal control regions and in the essential form of the optimal control laws. Therefore, the author hopes that Professor Desoer and his associates will consider extending the general theory to include the class of on-off systems.