Paper

Exact linearization of third-order systems with pulse-width-modulation-type inputs

Masayasu Suzuki¹ᵃ) and Mitsuo Hirata¹

¹ Graduate School of Engineering, Utsunomiya University
7-1-2 Yoto Utsunomiya, Tochigi 321-8585, Japan

ᵃ) ma-suzuki@cc.utsunomiya-u.ac.jp

Received July 31, 2017; Revised November 25, 2017; Published April 1, 2018

Abstract: An exact linearization method using static input transformations is proposed for third-order continuous-time systems with pulse-width-modulation-type inputs whose eigenvalues are real and distinct. In this method, two rectangular waves are located in each control interval, and their rising/falling timings are treated as the control parameters. This paper demonstrates that the image of the nonlinear input map contains the image of a suitable linear map, which means that some input transformations exist for linearization.

Key Words: pulse-width-modulation, nonlinearity, exact linearization, input transformation

1. Introduction

In practical situations concerning the analysis and design of control systems, we are often forced by technical/financial factors to implement actuators whose output values are limited to finite sets. For instance, the pulse-width-modulation (PWM) scheme is widely adopted for controlling electric motors, and contributes to a lower power consumption. Moreover, in controlling the flux of a gas/liquid, because it is difficult to accurately manipulate the opening of valves, one often makes the opening patterns finite in order to increase the reliability of a device. Since treating such systems with discrete events is difficult in general, it is a common practice to regard them as continuous-valued systems in some exact or approximate manner.

In this paper, we treat continuous-time linear time-invariant (LTI) systems with PWM-type inputs. In general a PWM input is defined as a time series where at most one rectangular wave is assigned on each control interval, and the ratio of the width of each rectangular wave to the control period, which is called the duty ratio, is adjustable. The variation of parameters leads to a discrete-time evolution equation describing the transition of the state of the continuous-time system on the sampling instants. The control input for such a discrete-time system is the sequence of the duty ratios, which means that we derive a system with a continuous-valued input. Nevertheless, because the input term is given as a nonlinear map for the duty ratio, it is still difficult to treat such a system. In fact, it is known that controlled PWM input systems can produce rich phenomena such as periodic and chaotic solutions. See, for example, [1] and references therein.

In the field of control theory, the stability of the zero solutions of feedback systems involving PWM input generators has been systematically discussed [2–7]. Meanwhile, for improving the control performance such as accomplishing high-speed positioning and path following, it is effective to apply
feedforward control. To design a valid feedforward input, an accurate model should be prepared, and moreover the input design will become tractable if the model for the input design is linear. In the control application field, some researchers have focused on linear approximations of the input terms of PWM systems [8, 9]. The study [8] mentions that for a PWM input system in which every rectangular wave is arranged at the center of the sampling period, the system linearized around 0% of the duty ratio agrees with the original nonlinear system with second-order accuracy. However, if the absolute values of the eigenvalues of the control object (more precisely, the products of these with the sampling period) become large, then so does the approximation error.

The present authors have proposed an exact linearization method for a class of second-order LTI systems with PWM-type inputs, where not only the duty ratio but also the location of the pulse wave is variable in the control interval [10]. To the best of the authors’ knowledge, there has been no research regarding the exact linearization prior to this work. However, extending the method in [10] such that it can be applied to higher-order systems is not straightforward, because the explicit construction of the solution presented in that paper appears to be difficult for higher-order cases.

This paper treats third-order LTI systems whose eigenvalues are real and distinct. We show that such systems with PWM-type inputs can be exactly linearized by an input transformation. To linearize the nonlinear input term, one requires more than three degrees-of-freedom for control inputs. To this end, two rectangular waves are placed in each control interval, and their widths and locations are made variable. This setup means that we derive four manipulable parameters; that is, the four timings of the rising/falling edges of the pulse waves. However, the first rising timing or second falling one can be fixed, as explained later. To confirm that the exact linearization is possible with this approach, we investigate the image of the nonlinear input map. In fact, the image is convex and contains the image of a desired linear map. The proposed input transformation is static, and it can be calculated in advance if the eigenvalues of the control object are known. Numerical examples for the input transformation will be presented in the last section.

Our study is essentially based on mathematical models, and tackles the development of theory and applications, which is in line with the aims of the FISRT Aihara Project. The study [11] constitutes a preliminary version of this paper. The proofs in the present paper have been presented more precisely, and some numerical simulations have been added.

2. Problem formulation and main result

Consider a third-order continuous-time LTI system

\[ \dot{x} = Ax + Bu, \]  

where \( x(t) \in \mathbb{R}^3 \) is the state of the system, \( u(t) \in \mathbb{R}^1 \) is the input, and \( A, B \) are real matrices. We assume the following:

1. The system matrix \( A \) has three distinct non-zero real eigenvalues \( \{\lambda_i\}_{i=1,2,3} \).
2. The input \( u \) consists of at most two rectangular pulses of height \( w^0 \) on each sampling interval of length \( T \). For a positive integer \( k \), we denote the locations of the rectangular pulses on \( [kT, (k+1)T] \) by ordered real numbers \( a_k, b_k, c_k, d_k \in [0,1] \) as \( [(a_k + k)T, (b_k + k)T] \) and \( [(c_k + k)T, (d_k + k)T] \). That is, the input \( u \) satisfies

\[
  u(t) = \begin{cases} 
    w_k, & t/T - k \in [a_k, b_k] \cup [c_k, d_k], \quad k = 0, 1, 2, \ldots, \\
    0, & \text{otherwise},
  \end{cases}
\]

\[
  w_k \in \{w^0, -w^0\}, \quad k = 0, 1, 2, \ldots,
\]

\[
  0 \leq a_k \leq b_k \leq c_k \leq d_k \leq 1, \quad k = 0, 1, 2, \ldots.
\]

We refer the reader to Fig. 1.

We consider the behavior of the system (1) on the sampling instants \( t = kT, k = 0, 1, 2, \ldots \). The transitions of the states \( \{x(kT)\} \) can be calculated through the variation of the parameters as follows:
Theorem 1. Through this static nonlinear transformation. The following constitutes the main result of this paper. That is, the state transitions intended for the linear system (5) can be realized where (3) and (4) with \( L \) then there exist unique smooth functions, \( \bar{a}, \bar{b}, \bar{c}, \bar{d} \):\([0,1]\) satisfying (3) and (4) with \( \bar{v} \) for any \((a,b,c,d)\) for each of the inputs \(v_k\) determined for (5), the behavior of the system (1) on the sampling points is the same as that of the system (5). That is, the state transitions intended for the linear system (5) can be realized through this static nonlinear transformation. The following constitutes the main result of this paper.

**Theorem 1.** For an arbitrarily chosen tuple \((a^*,b^*,c^*,d^*)\), let

\[
L = I (a^*, b^*, c^*, d^*). 
\]

Then there exist functions \( \bar{a}, \bar{b}, \bar{c}, \bar{d} : [0,1] \rightarrow [0,1] \) satisfying (3) and (4). In addition, suppose that \((A,B)\) is controllable. If the first rising timing \( a \) is fixed to 0 (or the last falling timing \( d \) is fixed to 1), then there exist unique smooth functions, \( \bar{b}, \bar{c}, \bar{d} : [0,1] \rightarrow [0,1] \) (or \( \bar{a}, \bar{b}, \bar{c} : [0,1] \rightarrow [0,1] \)) satisfying (3) and (4) with \( L = I (0, b^*, c^*, d^*) \) for any \((b^*,c^*,d^*)\) (or \( L = I (a^*, b^*, c^*, 1) \) for any \((a^*,b^*,c^*)\), respectively).

Consider a region

\[
\mathcal{D} = \{(a,b,c,d) \in \mathbb{R}^4 \mid 0 \leq a \leq b \leq c \leq d \leq 1\}. 
\]

The above result stems from the fact that the image of \( \mathcal{D} \) under \( I \) is convex. In the following section, we will confirm this property.

**Remark 2.** The latter part of Theorem 1 assumes that either of the first rising timing or last falling one is fixed on the sampling instant. This is to guarantee the uniqueness of the inverse image of each point in \( I(\mathcal{D}) \). Since there are three equations in (3) with four unknown variables, one variable must be eliminated for uniqueness. We will confirm in Section 3.4 that it is sufficient to fix \( a \) to 0 or \( d \) to 1. Even if we fix \( a \) to some positive number \( \hat{a} > 0 \), the latter part of Theorem 1 remains valid by replacing the value \( L = I (0, b^*, c^*, d^*) \) with \( L = I (\hat{a}, b^*, c^*, d^*) \). However, an inclusion relation

\[
\{I (\hat{a}, b, c, d) \mid \hat{a} \leq b \leq c \leq d \leq 1\} \subseteq \{I (0, b, c, d) \mid 0 \leq b \leq c \leq d \leq 1\}
\]
holds, which implies that the linearization when fixing \( a = \hat{a} \) is accomplished whenever we fix \( a = 0 \). Because fixing \( a \) to a value other than 0 only restricts the set of the input vectors of (5) to a small subset of \( I(D) \), this makes no sense. We can say the same for fixing \( d \) to a value other than 1.

When \( b \) is fixed to some value \( \hat{b} \), the image \( \{ I(\hat{a}, \hat{b}, c, d) \mid 0 \leq a \leq \hat{b} \leq c \leq d \leq 1 \} \) is a non-convex set (the yellow filled tube in Fig. 7). This leads to a failure to have solutions for (3)–(4) for some \( L \) even though \( L \) is chosen to be of the form \( L = I(\hat{a}^*, \hat{b}^*, c^*, d^*) \). From this point of view, we do not consider fixing \( b \) in this paper. The same can be said for fixing \( c \).

**Remark 3.** Although it is assumed in this paper that the system matrix \( A \) has three distinct non-zero real eigenvalues, we expect that this assumption can be relaxed. In fact, for second-order systems with PWM-type inputs where each control interval has a single pulse with a pair of variable rising/falling timings \((a, b)\), one can confirm the following properties regarding the two-dimensional image \( I_{2D} \) corresponding to \( I(D) \) [12]:

- \( I_{2D} \) is convex as long as the eigenvalues are real. If the two eigenvalues are distinct, or if the geometric multiplicity is not two even though they are the same, then \( I_{2D} \) has a nonempty interior that is homeomorphic to \( \Delta := \{(a, b) \mid 0 < a < b < 1\} \).

- Suppose that the eigenvalues are complex conjugate to one another. If the absolute value of the imaginary parts is less than the Nyquist angular frequency \( \pi/T \), then \( I_{2D} \) is a convex body whose interior is homeomorphic to \( \Delta \). Otherwise, \( I_{2D} \) becomes a star-shaped set with respect to the origin, owing to a kind of folding effect. In the latter case, while the homeomorphic property is lost, the linearization can be still accomplished.

Here, we suppose that \((A, B)\) is controllable. The proofs are essentially similar among the various classes of the eigenvalues, but there are technically different details. Thus, such second-order systems are always linearizable, while the richness of the obtainable linear systems and uniqueness of the transformation depend on the eigenvalues. From these facts, the authors are positive regarding a conjecture that third-order systems are also always linearizable. Extending the result of this paper to wider classes of systems, including higher-dimensional systems, will be a direction for future work.

### 3. The image \( I(D) \)

Let \( V \) be a diagonalization matrix of \( A \), i.e., \( A = V \text{diag}\{\lambda_i\} V^{-1} \), and denote \( V^{-1} B = \begin{bmatrix} \tilde{b}_1 & \tilde{b}_2 & \tilde{b}_3 \end{bmatrix}^T \). In addition, let \( \nu_i = -\lambda_i T, i = 1, 2, 3, \) and introduce a matrix \( K = \text{diag}\{1/k_i\} \) for sake of a scale adjustment, where \( k_i = 1/(1 - e^{\nu_i}), i = 1, 2, 3 \). Then, we have that

\[
I(a, b, c, d) = A^{-1} e^{AT} V \text{diag}\{\tilde{b}_i\} K m(a, b, c, d),
\]

where

\[
m(a, b, c, d) = \begin{bmatrix}
    k_1 (e^{\nu_1 a} - e^{\nu_1 b} + e^{\nu_1 c} - e^{\nu_1 d}) \\
    k_2 (e^{\nu_2 a} - e^{\nu_2 b} + e^{\nu_2 c} - e^{\nu_2 d}) \\
    k_3 (e^{\nu_3 a} - e^{\nu_3 b} + e^{\nu_3 c} - e^{\nu_3 d})
\end{bmatrix}
\]

We investigate the image \( m(D) \) instead of \( I(D) \). In fact, it turns out that if \( m(D) \) is convex, then so is \( I(D) \). Note that the map \( m \) is determined only by the eigenvalues of \( A \) and the sampling period \( T \).

The image \( m(D) \) consists of two smooth surfaces \( m(\Delta^1) \) and \( m(\Delta^2) \), and the region surrounded by these surfaces, where

\[
\Delta^1 = \{(a, b, 1, 1) \in D \mid 0 \leq a \leq b \leq 1\},
\]

\[
\Delta^2 = \{(0, b, c, 1) \in D \mid 0 \leq b \leq c \leq 1\}.
\]

For the sake of simple notation, we define the following map:
Fig. 2. The image of $D$ under $m$, $m(D) \subset \mathbb{R}^3$.

Fig. 3. The images of $\Delta^1$ and $\Delta^2$ under $m$ when $\Lambda < 0$. For each triangle region, the two sides are mapped to the boundary of the image, and the diagonal line is mapped to a point ($\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$) for $\Delta^1$ and $\Delta^2$, respectively). The arrows on the surfaces indicate the directions in which the parameters $a$, $b$, and $c$ increase.

$$S(\alpha, \beta) = \left[ k_1 \left( e^{\nu_1 \alpha} - e^{\nu_1 \beta} \right) \quad k_2 \left( e^{\nu_2 \alpha} - e^{\nu_2 \beta} \right) \quad k_3 \left( e^{\nu_3 \alpha} - e^{\nu_3 \beta} \right) \right]^T.$$ 

Then the surfaces can be represented as

$$m(\Delta^1) = \{ S^1(a, b) \in \mathbb{R}^3 \mid 0 \leq a \leq b \leq 1 \},$$

$$m(\Delta^2) = \{ S^2(b, c) \in \mathbb{R}^3 \mid 0 \leq b \leq c \leq 1 \},$$

where $S^1(a, b) = S(a, b)$ and $S^2(b, c) = 1_{3 \times 1} - S(b, c)$. Denote the surfaces by $S^1$ and $S^2$, which are the same as the maps describing themselves. (We also call the map $S$ “surface $S$.”) Figure 2 depicts the image $m(D)$. Figure 3 illustrates the correspondences between the regions $\{ \Delta^i \}_{i=1,2}$ and surfaces $\{ S^i \}_{i=1,2}$.

3.1 Curvatures of the surfaces

To investigate the curvature of the surface $S$, we calculate its second fundamental form:

$$\Pi := P dd\alpha + 2Q d\alpha d\beta + R d\beta d\beta,$$

where

$$P = S_{\alpha\alpha} \cdot n, \quad Q = S_{\alpha\beta} \cdot n, \quad R = S_{\beta\beta} \cdot n.$$
The symbols with the subscripts \( \alpha \) and \( \beta \) denote the corresponding partial derivatives, e.g., \( S_\alpha = \partial S/\partial \alpha \) and \( S_{\alpha\alpha} = \partial^2 S/\partial \alpha^2 \). The symbol \( n \) denotes the normal vector to the surface at the point \( S(\alpha, \beta) \), where \( \alpha \neq \beta \). The normal vector \( n \) is defined by

\[
 n := \begin{cases} 
  S_\alpha \times S_\beta, & \text{when } \Lambda < 0, \\
  S_\beta \times S_\alpha, & \text{when } \Lambda > 0, 
\end{cases} 
\]

(7)

The operators "\( \times \)" and "\( \times \)" denote the inner and exterior products, respectively. Since the tangent vectors \( S_\alpha(\alpha, \beta) \) and \( S_\beta(\alpha, \beta) \) are linearly independent as long as \( \alpha \neq \beta \), for every point on the surface other than the origin \([0 \ 0 \ 0]^T\), the tangent plane and normal vector can be defined. We are only interested in the sign of \( \Pi \), and hence the length of \( n \) is not normalized. In addition, the reason for changing the definition of \( n \) depending on \( \text{sgn}(\Lambda) \) is to unify the directions of the normal vectors with the external side of the image. As shown below, \( P \) and \( R \) are always negative under this definition of the normal vectors.

**Lemma 4.** The second fundamental form \( \Pi \) is negative definite. That is, the surface \( \{S(\alpha, \beta) \mid \alpha < \beta\} \) is convex to the direction of the normal vector at each point \( S(\alpha, \beta) \).

**Proof.** We have immediately that \( Q = 0 \). \( P \) and \( Q \) can be calculated as follows:

\[
 P = -\text{sgn}(\Lambda)k_1k_2k_3\nu_1\nu_2\nu_3e^{(\nu_1+\nu_2+\nu_3)\alpha} \det \begin{bmatrix} 1 & 1 & 1 \\ \nu_1 & \nu_2 & \nu_3 \\ e^{\nu_1(\beta-\alpha)} & e^{\nu_2(\beta-\alpha)} & e^{\nu_3(\beta-\alpha)} \end{bmatrix},
\]

\[
 R = -\text{sgn}(\Lambda)k_1k_2k_3\nu_1\nu_2\nu_3e^{(\nu_1+\nu_2+\nu_3)\beta} \det \begin{bmatrix} 1 & 1 & 1 \\ \nu_1 & \nu_2 & \nu_3 \\ e^{\nu_1(\alpha-\beta)} & e^{\nu_2(\alpha-\beta)} & e^{\nu_3(\alpha-\beta)} \end{bmatrix}.
\]

Since the functions \( \nu \mapsto e^{\nu(\beta-\alpha)} \) and \( \nu \mapsto e^{\nu(\alpha-\beta)} \) are strictly convex, it turns out from Fact 10 in Appendix that the sign of the above determinant is opposite to that of \( \Lambda \). Taking into account that \( k_1k_2k_3\nu_1\nu_2\nu_3 < 0 \), we have that \( P < 0 \) and \( R < 0 \). Thus, the second fundamental form is negative definite. \( \square \)

For the surfaces \( S^1 \) and \( S^2 \), define the normal vectors at the points \( S^1(a, b) \) and \( S^2(b, c) \) as follows:

\[
 n^1(a, b) := \text{sgn}(\Lambda)S^1_b(a, b) \times S^1_a(a, b),
\]

\[
 n^2(b, c) := \text{sgn}(\Lambda)S^2_c(b, c) \times S^2_b(b, c),
\]

where \( S^1_b(a, b) = S_\alpha(a, b) \), \( S^1_a(a, b) = S_\beta(a, b) \), \( S^2_b(b, c) = -S_\alpha(b, c) \), and \( S^2_c(b, c) = -S_\beta(b, c) \).

**Lemma 5.** Consider the normal vectors on the surfaces \( S^1 \) and \( S^2 \). The signs of their elements on \( S^1 \) are consistent, and so are those on \( S^2 \). Moreover, the angle formed by a normal vector on \( S^1 \) and one on \( S^2 \) is always in \([-\pi, -\pi/2) \cup (\pi/2, \pi]\).

**Proof.** From the definition of \( n^1 \), we derive that

\[
 n^1(a, b) = \text{sgn}(\Lambda) \begin{bmatrix} k_2k_3k_1\nu_1\nu_2\nu_3(e^{\nu_1a+b+\nu_2b} - e^{\nu_3a+\nu_2b}) \\ k_3k_1k_2\nu_1\nu_3\nu_2(e^{\nu_1a+b+\nu_3b} - e^{\nu_2a+\nu_3b}) \\ k_1k_2k_3\nu_1\nu_2\nu_3(e^{\nu_1a+\nu_2b} - e^{\nu_3a+\nu_2b}) \end{bmatrix}.
\]

For distinct \( i, j \in \{1, 2, 3\} \), the equation \( e^{\nu_i a+b} - e^{\nu_j a+b} = 0 \) is equivalent to \((\nu_i - \nu_j)(a - b) = 0\). Since \( \nu_i \neq \nu_j \), the term \( e^{\nu_i a+b} - e^{\nu_j a+b} \) does not vanish as long as \( a < b \), which means that the sign does not change. Since it holds that \( n^2(b, c) = -n^1(b, c) \), the signs of the elements of the normal vectors on \( S^2 \) are also consistent, and moreover they are opposite to those on \( S^1 \). Therefore, for any \((a^1, b^1) \in \Delta^1 \) and \((b^2, c^2) \in \Delta^2 \), it holds that

\[
 n^1(a^1, b^1) \cdot n^2(b^2, c^2) = -n^1(a^1, b^1) \cdot n^1(b^2, c^2) < 0,
\]

which implies that the angle formed by \( n^1 \) and \( n^2 \) is in \([-\pi, -\pi/2) \cup (\pi/2, \pi]\). \( \square \)
From Lemmas 4 and 5, we find that the surfaces $S^1$ and $S^2$ are convex toward the normal vectors on each surface, where the normal vectors on the same surface trend in almost the same direction, which are opposite to the directions of those on the other surface. Recall that $S^1$ and $S^2$ have a common boundary. Thus, the geometry of the set $S^1 \cup S^2$ looks like a clam.

### 3.2 Angles of the surfaces at the boundary

Next, we investigate how the two surfaces $S^1$ and $S^2$ combine at their common boundary. The common boundary consists of two sub-boundaries $B^u$ and $B^d$, which are defined below. For arbitrary $b_* \in (0,1)$, it holds that $S^1(0, b_*) = S^2(b_*, 1)$, and the value increases up to $[1 \ 1]^\top$ from the origin as $b_*$ increases. We denote this part of the boundary by $B^u$. In the same manner, since it holds for arbitrary $a_* \in (0,1)$ that $S^1(a_*, 1) = S^2(0, a_*)$ decreases down to the origin as $a_*$ increases, we denote this sub-boundary by $B^d$. Note that for each surface, tangent planes can be naturally defined, even on the boundary by extending the domain of $S$. We denote the tangent planes on the boundary as follows:

- **Tangent planes on the sub-boundary $B^u$:**
  
  $$P_{b_*}^{1u} := \text{span} \{ S^1_0(0, b_*), S^2_0(0, b_*) \} \quad \text{on} \ S^1,$$
  
  $$P_{b_*}^{2u} := \text{span} \{ S^2_0(b_*, 1), S^2_0(b_*, 1) \} \quad \text{on} \ S^2.$$

- **Tangent planes on the sub-boundary $B^d$:**
  
  $$P_{a_*}^{1d} := \text{span} \{ S^1_0(a_*, 1), S^1_0(a_*, 1) \} \quad \text{on} \ S^1,$$
  
  $$P_{a_*}^{2d} := \text{span} \{ S^2_0(0, a_*), S^2_0(0, a_*) \} \quad \text{on} \ S^2.$$

**Lemma 6.** For arbitrary $b_* \in (0,1)$, the dihedral angle formed by $P_{b_*}^{1u}$ and $P_{b_*}^{2u}$ is in $(\pi/2, \pi)$. Meanwhile, for arbitrary $a_* \in (0,1)$, the dihedral angle formed by $P_{a_*}^{1d}$ and $P_{a_*}^{2d}$ is in $(\pi/2, \pi)$.

**Proof.** Consider the sub-boundary $B^u$ (see Fig. 4). Note that the line of intersection between the tangent planes $P_{b_*}^{1u}$ and $P_{b_*}^{2u}$ includes the tangent vector $S^1_0(0, b_*) = S^2_0(b_*, 1)$, which we denote by $t_b$. The dihedral angle is defined as the angle formed by the normal vectors of the two planes. If these two normal vectors, denoted simply by $n^1$ and $n^2$, and the tangent vector $t_b$ compose a right-handed system in the order of $(n^2, n^1, t_b)$ (or $(n^1, n^2, t_b)$) for $\Lambda < 0$ (for $\Lambda > 0$, respectively), then the dihedral angle is in $(0, \pi)$. This condition is equivalent to the relation $\text{sgn}(\Lambda)(t_b \cdot S^1_a \times S^2_c) > 0$ since

$$n^1 \times n^2 = (t_b \times S^1_a) \times (t_b \times S^2_c) = (t_b \cdot S^1_a \times S^2_c) t_b.$$

We can derive by a simple calculation that

$$t_b \cdot S^1_a \times S^2_c = k_1 k_2 k_3 \nu_1 \nu_2 \nu_3 \det \begin{bmatrix} 1 & 1 & 1 \\ e^{\nu_1 b_*} & e^{\nu_2 b_*} & e^{\nu_3 b_*} \\ e^{\nu_1} & e^{\nu_2} & e^{\nu_3} \end{bmatrix}.$$

The matrix above is called the generalized Vandermonde matrix, and the sign of its determinant coincides with $-\text{sgn}(\Lambda)$ for any $b_* \in (0,1)$ (see Fact 11 in Appendix). Since $k_i \nu_i < 0$, $i = 1, 2, 3$, it holds that $\text{sgn}(\Lambda)(t_b \cdot S^1_a \times S^2_c) > 0$. In the same manner, one can confirm the result for the sub-boundary $B^d$. \qed

This lemma says that while $S^1$ and $S^2$ stick together on the common boundary, there is no valley close to the boundary when viewed from the exterior of $S^1 \cup S^2$. See the left figure in Fig. 5, which explains this condition.
3.3 Inverse images of the interior

Finally, we confirm that \( m(D) \) has no hollow, or more strictly that for each element in the region surrounded by \( S^1 \) and \( S^2 \) there exist inverse images \( (a, b, c, d) \in D \). Choose \( b_s \in [0, 1] \) arbitrarily, and let \( a = 0 \) and \( b = b_s \). Then, consider the following two-dimensional surfaces parametrized by \( (c, d) \):

\[
U(b_s) := \{ m(0, b_s, c, d) \mid b_s \leq c \leq d \leq 1 \}.
\]

Figure 6 shows a transition of the surfaces \( U(b_s) \) when \( b_s \) varies from 0 to 1. Each surface has two boundaries \( \{ m(0, b_s, c, 1) \mid c \in [b_s, 1] \} \) and \( \{ m(0, b_s, b_s, d) \mid d \in [b_s, 1] \} \). The former is contained in \( S^2 \) and the union over \( b_s \in [0, 1] \) covers \( S^1 \). The latter is always contained in \( B^u \). In addition, \( U(0) \) is equal to \( S^1 \) since \( \{ m(0, 0, c, d) \mid 0 \leq c \leq d \leq 1 \} = \{ m(a, b, 0, 0) \mid 0 \leq a \leq b \leq 1 \} \).

Thus it appears that the family of \( U(b_s) \) fills the region surrounded by \( S^1 \) and \( S^2 \). The following lemma justifies this conjecture.

**Lemma 7.** Each surface \( U(b_s), b_s \in [0, 1] \) is contained in the region surrounded by \( S^1 \) and \( S^2 \). Furthermore, the union \( \bigcup_{b_s \in [0, 1]} U(b_s) \) coincides with this region.

**Proof.** Fix \( b_s \in [0, 1] \), and consider the following two-dimensional surfaces parametrized by \( (c, d) \):

\[
V_{b_s}(a_t) := \{ m(a_t, b_s, c, d) \mid b_s \leq c \leq d \leq 1 \},
\]

where \( a_t \in [0, b_s] \). Figure 7(a) shows the transition of \( V_{b_s}(a_t) \) when \( a_t \) varies from \( b_s \) to 0. All surfaces \( V_{b_s}(a_t), a_t \in [0, b_s] \) are congruent, and their orientations are the same. Note that \( V_{b_s}(0) = U(b_s) \), and
This figure shows the surfaces $U(0)$ and \{ $U(b_s)$ \}$_{i=1}^4$ where $0 < s_1 < s_2 < s_3 < s_4 < 1$. Note that $U(0) = S^1$. (b) Each surface is surrounded by two boundaries: One is the curve \{ $m(0, b_s, c, 1)$ \}$_{c \in [b_s, 1]}$, and the other is \{ $m(0, b_s, b_s, d)$ \}$_{d \in [b_s, 1]}$.

Furthermore, they are swept in the yellow tube $T^b$ (or red tube $T^u$) as $a_t (b_t)$ varies. Moreover $V_{b_s}(b_s)$ is a subset of $S^1$. Since the surfaces $V_{b_s}(a_t)$ are swept along the boundary sub-curve 
\{ $S^1(a_t, 1)$ \}$_{a_t \in [0, b_s]}$ \subset $B^d$,

the direction of movement of the surface $V_{b_s}(a_t)$ in the sweeping procedure is given by the tangent vector of this curve as follows:

$$-S_{b_s}(a_t, 1) = - \begin{bmatrix} k_1 \nu_1 e^{\nu_1 a_t} & k_2 \nu_2 e^{\nu_2 a_t} & k_3 \nu_3 e^{\nu_3 a_t} \end{bmatrix}^T.$$

Denote this tangent vector by $\xi(a_t)$.

Consider the inner product of the normal vector on $V_{b_s}(a_t)$ and the tangent vector $\xi(a_t)$. Since $V_{b_s}(a_t)$ is congruent with $V_{b_s}(b_s) = \{ S^1(a, b) \mid b_s \leq a \leq b \leq 1 \}$, one can use the normal vectors on $V_{b_s}(b_s)$ instead of those on $V_{b_s}(a_t)$. Then, we have that

$$ n^1(a, b) \cdot \xi(a_t) = -\text{sgn}(\Lambda) k_1 k_2 k_3 \nu_1 \nu_2 \nu_3 \det \begin{bmatrix} e^{\nu_1 a_t} & e^{\nu_2 a_t} & e^{\nu_3 a_t} \\ e^{\nu_1 a} & e^{\nu_2 a} & e^{\nu_3 a} \\ e^{\nu_1 b} & e^{\nu_2 b} & e^{\nu_3 b} \end{bmatrix} < 0,$$

where the relation $a_t \leq a \leq b$ and Fact 11 are used. Thus, the angle formed by any normal vector with the tangent vectors is greater than $\pi/2$. This means that in the sweeping procedure, $V_{b_s}(a_t)$
does not protrude from the tube $T^d$ surrounded by curves congruent with $B^d$, defined by

$$T^d = T^d_1 \cup T^d_2,$$

$$T^d_1 = \{m(a, b_s, b_s, d) \mid 0 \leq a \leq b_s \leq d \leq 1\}, \quad T^d_2 = \{m(a, b_s, c, 1) \mid 0 \leq a \leq b_s \leq c \leq 1\}.$$  

The smooth surfaces $T^d_1$ (the bottom-half part of the tube) and $T^d_2$ (the upper-half) share points only on their common boundaries, and moreover $T^d \cap S^1 = T^d_1$. Therefore, it turns out that the tube $T^d$ sticks close to $S^1$ (see Fig. 7(a)), and hence $U(b_s)$ does not intersect with $S^1$ other than on the common boundary.

Next, we consider a family of surfaces

$$W_{b_s}(s) := \{m(0, b_s, c, d) \mid b_s \leq c \leq d \leq 1\}, \quad b_s \in [0, b_s],$$

and the following tube $T^u$ surrounded by curves congruent with a boundary sub-curve on $B^u$:

$$T^u = \{m(0, b_s, c, d) \mid 0 \leq b \leq b_s \leq c \leq 1\} \cup \{m(0, b_s, d) \mid 0 \leq b \leq b_s \leq d \leq 1\}.$$  

Note that $W_{b_s}(s) = U(b_s)$, $W_{b_s}(0) = V_{b_s}(s)$, and the tube $T^u$ sticks close to $S^2$. Then, it turns out through the same calculation as above that the family of surfaces is included in the tube (see Fig. 6(b)), which means that $U(b_s)$ and $S^2$ do not intersect other than on the boundary of $U(b_s)$. This implies that $U(b_s)$ in the contained the region between $S^1$ and $S^2$.

The sub-boundary of $U(b_s)$, $\{m(0, b_s, d) \mid d \in [b_s, 1]\}$, is monotonic with respect to $b_s$. That is, for $s$ and $\tilde{s}$ that satisfy $0 < s < \tilde{s} < 1$, it holds that $(U(b_s) \cap B^s) \subset (U(b_{\tilde{s}}) \cap B^s)$. In addition to this nesting structure, considering the fact that the family of $U(b_s)$ are smoothly swept inside of the region surrounded by $S^1$ and $S^2$, we find that $\bigcup_{b_s \in [0, 1]}$ does not have any hollow.

From the above discussions, we obtain the following result.

**Theorem 8.** The image $m(D)$ is convex.

**Proof.** Fix $b_s \in [0, 1]$ arbitrarily, and consider sets $U_a(b_s) := \{m(a, b_s, c, d) \mid b_s \leq c \leq d \leq 1\}$ for $a \in [0, b_s]$. Since it holds that

$$m(a, b_s, c, d) = m(0, b_s, c, d) - S(0, a), \quad (a, b_s, c, d) \in D,$$

we find that $U_a(b_s)$ is a parallel-translated set of $U(b_s)$ along $B^u$. In addition, by the same reasoning as in Lemma 7, $U_a(b_s)$ is contained in the region surrounded by $S^1$ and $S^2$. Therefore, for all elements $(a, b_s, c, d)$ in $D$ the images are in the region surrounded by $S^1$ and $S^2$, which implies that $m(D)$ coincides with the region surrounded by $S^1$ and $S^2$.

Note that the closed surface $S^1 \cup S^2$ constitutes the whole of the boundary of $m(D)$, and moreover $m(D) \setminus (S^1 \cup S^2)$ is a connected open set. In addition, we find from the arguments in Sections 3.1 and 3.2 that $m(D) \setminus (S^1 \cup S^2)$ has a local support plane at every boundary point. Therefore, by the Erhard Schmidt theorem (see Fact 12 in Appendix) it follows that $m(D) \setminus (S^1 \cup S^2)$ is convex. Since $m(D)$ is the closure of $m(D) \setminus (S^1 \cup S^2)$ and the closure of a convex set is convex, we can conclude that $m(D)$ is convex.

One can immediately derive the first argument in Theorem 1 from Theorem 8.

### 3.4 Uniqueness of the transformation law

There are four degrees of freedom for the mapping $m$, while its range is in $\mathbb{R}^3$. Hence, it is clear that for each $v \in [0, 1]$ there exist multiple solutions of $m(a, b, c, d) = Lv$. However, according to Sec. 3.3, the family $\{U(b_s)\}_{b_s \in [0, 1]}$, where the first rising timing $a$ is fixed to zero, can cover the image $m(D)$. Let $m_0(b, c, d) := m(0, b, c, d)$, and consider the following equation:

$$\begin{cases}
  m_0(b, c, d) = Mv, \\
  0 \leq b \leq c \leq d \leq 1,
\end{cases}$$

where $M$ is given in advance as $M = m_0(b^*, c^*, d^*)$ for some $(0, b^*, c^*, d^*) \in D$. Then, the following holds.
Proposition 9. For each \( v \in (0, 1) \), the Eq. (8) has a unique solution \((b, c, d)\). Moreover, this solution is smooth with respect to \( v \).

Proof. For distinct \( b_\ell \) and \( b'\) in \([0,1]\), we have that \( U(b_\ell) \cap U(b_\ell') \subset B^u \). Therefore, the line segment \( Mv \) only intersects one surface in \( \{U_{b_\ell}\}_{b_\ell \in (0,1)} \) at each \( v \in (0,1) \).

Consider a function \( F(b,c,d;v) = m_0(b,c,d) - Mv \). Note that the Jacobian \( D_{(b,c,d)} F \) is the generalized Vandermonde matrix, and is nonsingular. Suppose that for some \((b',c',d')\) and \(v'\), it holds that \( F(b',c',d';v') = 0 \). From the implicit function theorem, there exist an open interval \( J \subset (0,1) \) around \( v' \) and a unique \( C^\infty \)-class function \( g : J \to \{(b,c,d) \mid 0 < b < c < d < 1\} \) such that \( g(v') = (b',c',d') \) and \( F(g(v);v) = 0 \) hold for \( v \in J \).

Note that when \((A,B)\) is controllable, \( \text{diag}\{b_\ell\} \) is nonsingular, and hence the sets \( \{I(0,b,c,d) \mid 0 < b < c < d < 1\} \) and \( \{m_0(b,c,d) \mid 0 < b < c < d < 1\} \) are diffeomorphic. From Proposition 9, the second argument for fixing \( a \) to 0 in Theorem 1 can be immediately derived.

To guarantee the uniqueness, one can set \( d = 1 \) instead of fixing \( a \) for the following reason: (i) the set \( \{I(a,b,c,1) \mid 0 < a \leq b \leq c \leq 1\} \) is convex, which can be immediately concluded from the relation \( I(a,b,c,1) = I(0,0,0,1) - I(0,a,b,c) \); (ii) the above set includes the line segments described as \( \{I(a^*,b^*,c^*,1) v \mid v \in [0,1]\} \), where \( a^*,b^*, \) and \( c^* \) are arbitrary; (iii) by similar reasoning as in the case that \( a = 0 \), we can confirm that the solution of \( I(a,b,c,1) = I(a^*,b^*,c^*,1) v \) is unique. Thus, the second argument for fixing \( d \) to 1 in Theorem 1 is also true.

4. Numerical simulations

In this section, we confirm through numerical simulations that the proposed method is efficient compared with a conventional linear approximation method. Consider the system (1) with the following coefficient matrices:

\[
A = \begin{bmatrix}
-4 & 0 & 0 \\
0 & -8 & 0 \\
0 & 0 & -10
\end{bmatrix}, \quad B = \begin{bmatrix}
1 \\
1 \\
1
\end{bmatrix}.
\]

We apply PWM-type signals to (1) corresponding to a sinusoidal signal \( \sin \omega t \), \( \omega = 2\pi \times 0.01 \), where the height of the pulse is 1 (\( w^0 = 1 \)). Let \( T \) be the sampling period. Here, we set \( T = 1 \) or \( T = 2 \).

Method 1 We input pulse sequences to the continuous-time system (1), where each pulse is assigned at the center of each control interval. The manipulable variable is the width of the pulse, which is denoted as \( \delta_{k'} \). This situation appears in many applications, in particular when the so-called triangle comparison method is employed. Since the proposed method contains two pulses in control interval, we set the control period as \( T/2 \) to make the conditions fair. Then, the state transition obeys the following:

\[
x((k'+1)T/2) = e^{AT/2}x(k'T) + \int_{T/4-\delta_{k'}}^{T/4+\delta_{k'}} e^{A(T/2-\tau)} Bd\tau, \quad \delta_{k'} \in [0,T/2].
\]

By calculating the linear approximation system around \( \delta_{k'} = 0 \), we have that

\[
x_{k'+1} = e^{AT/2}x_{k'} + e^{AT/4}B\delta_{k'}.
\]

Method 2 We input pulse sequences to the continuous-time system (1), where the first rising timings \( \{a_k\} \) are always fixed to zero, and the other timings \( \{b_k\}, \{c_k\} \) and \( \{d_k\} \) are manipulable. Then the state transition obeys (2); that is,

\[
x((k+1)T) = e^{AT}x(kT) + w_k I(0,b_k,c_k,d_k).
\]

For (11), we consider the following linear system:

\[
x_{k+1} = e^{AT}x_k + Lv_k,
\]

214
where \( L = I(0, 0, 0, 1) \).

Let \( M = m_0(0, 0, 1) \). Although we could confirm the existence and uniqueness of the Eq. (8), it is difficult to solve analytically. Hence, we solve (8) numerically here. We reformulate the problem as

\[
\min_{b,c,d} \| m_0(b, c, d) - Mv \| \quad \text{subject to} \quad 0 < b < c < d < 1,
\]

for \( v \in (0, 1) \). Figure 8 presents the numerical solutions of the above optimization problem. We employ these solutions as the transformation law \( \bar{b}, \bar{c}, \bar{d} \), and determine the timings and sign of the pulses for (11) by \( b_k = \bar{b}(|v_k|), c_k = \bar{c}(|v_k|), d_k = \bar{d}(|v_k|) \), and \( w_k = w^0 \text{sgn}(v_k) \).

For (9) and (10) in the approximation method with centered pulses, the input \( \delta_k \) is given as \( \delta_{k'} = (T/2) \sin(\omega Tk'/2) \). Meanwhile, for (12) in the proposed method, the fictitious input \( v_k \) is given as \( v_k = \sin(\omega Tk) \), and \( (b_k, c_k, d_k) \) for (11) is determined against \( v_k \). Figure 9 shows the time responses of the systems (9)–(12). The solid lines (gray) and dashed lines (red) in the two left figures represent the responses of (9) and (10), respectively. We find that for Method 1, errors are presented between the system (9), which exactly captures the state transition on the sampling instants, and its approximation model (10). Moreover, the error becomes large when the sampling period increases. On the other hand, for the proposed Method 2 the responses of the nonlinear system (11) and linearized one (12) coincide, which can be seen from the right figures.
5. Conclusions
In this paper, we proposed an exact linearization method with static input transformations for third-order continuous-time LTI systems with PWM-type inputs whose eigenvalues are real and distinct. This method uses only information regarding the eigenvalues of the control object and the sampling period. In the proposed method, two rectangular waves are assigned in each control period, and their rising/falling timings are treated as the manipulable parameters. For this formulation, we investigated the image of the nonlinear input map, and found that it contains the image of a suitable linear map, and moreover, the first rising timing or second falling one can be fixed to the sampling point. We also presented a numerical example for the input transformation.

One of direction of our future work will be to extend the result of this paper to three-dimensional systems with complex eigenvalues, as well as higher-dimensional systems.

Acknowledgments
This work was partially supported by JSPS KAKENHI Grant Numbers JP17K14698 and JP16K14282.

Appendix

A. Some facts

Fact 10. Let $J$ be a finite or infinite interval, and let $f : J \rightarrow \mathbb{R}$. Then, the following statements are equivalent:

- $f$ is strictly convex.
- For all $x, y, z \in J$ such that $x < y < z$, it holds that

$$\det \begin{bmatrix} 1 & 1 & 1 \\ x & y & z \\ f(x) & f(y) & f(z) \end{bmatrix} > 0.$$ 

See page 21 of [13].

Fact 11. For real numbers $\{a_i\}_{i=1}^n$ and $\{p_k\}_{k=1}^n$ such that $0 < a_1 < a_2 < \cdots < a_n$, $p_1 < p_2 < \cdots < p_n$, consider a matrix $V = [a_i^{p_k}]_{i,k=1}^n$, which is called the generalized Vandermonde matrix. Then, it holds that

$$\det V > 0.$$ 

For $n = 3$, one can confirm this result from Fact 10. For a general $n$, see page 99 of [14].

Fact 12 (Erhard Schmidt Theorem). A connected open set in $n$-dimensional Euclidean space having a local support $(n-1)$-plane at every boundary point is convex.

See the appendix of [15].

References

[1] F. Angulo, E. Fossas, and G. Olivar, “Transition from periodicity to chaos in a PWM-controlled buck converter with ZAD strategy,” Int. J. Bifurcation and Chaos, vol. 15, no. 1, pp. 3245–3264, 2015.

[2] V.M. Kuntsevich and Y.N. Chekhovoi, “Fundamentals of non-linear control systems with pulse-frequency and pulse-width modulation,” Automatica, vol. 7, no. 1, pp. 73–81, 1971.

[3] L. Hou and A.N. Michel, “Stability analysis of pulse-width-modulated feedback systems,” Automatica, vol. 37, no. 9, pp. 1335–1349, 2001.
[4] S. Almér, U. Jönsson, C.Y. Kao, and J. Mari, “Stability analysis of a class of PWM systems,” IEEE Trans. Autom. Control, vol. 52, no. 6, pp. 1072–1078, 2007.
[5] H. Fujioka, C.Y. Kao, S. Almér, and U. Jönsson, “Robust tracking with $H_\infty$ performance for PWM systems,” Automatica, vol. 45, no. 8, pp. 1808–1818, 2009.
[6] F.R. Delfeld and G.J. Murphy, “Analysis of pulse-width-modulated control systems,” IRE Trans. Autom. Control, vol. 6, no. 3, pp. 283–292, 1961.
[7] S. Mariéthoz, S. Almér, M. Bája, A.G. Beccuti, D. Patino, A. Wernrud, J. Buisson, H. Cormerais, T. Geyer, H. Fujioka, U.T. Jönsson, C.Y. Kao, M. Morari, G. Papafotiou, A. Rantzer, and P. Riedinger, “Comparison of hybrid control techniques for buck and boost DC-DC converters,” IEEE Trans. Control Syst. Technol., vol. 18, no. 5, pp. 1126–1145, 2010.
[8] K.P. Gokhale, A. Kawamura, and R.G. Hoft, “Dead beat microprocessor control of PWM inverter for sinusoidal output waveform synthesis,” IEEE Trans. Ind. Appl., vol. IA-23, no. 5, pp. 901–910, 1987.
[9] H. Fujimoto, Y. Hori, and S. Kondo, “Perfect tracking control based on multirate feedforward control and applications to motion control and power electronics: A simple solution via transfer function approach,” Proc. Power Conversion Conference, pp. 196–201, 2002.
[10] M. Suzuki and M. Hirata, “Exact linearization of PWM-hold discrete-time systems using input transformation,” Proc. European Control Conference, pp. 446–451, 2015.
[11] M. Suzuki and M. Hirata, “Exact linearization of three-dimensional LTI systems with PWM inputs,” Proc. IFAC Conference on Analysis and Control of Chaotic Systems, pp. 267–272, 2015.
[12] M. Suzuki and M. Hirata, “Exact linearization of second-order PWM-type input systems: Star-shaped input-value sets” (in preparation).
[13] C. Niculescu and L.E. Persson, Convex Functions and Their Applications: A Contemporary Approach, Springer, New York, 2006.
[14] F.R. Gantmacher, The Theory of Matrices: Volume II, Chelsea Publishing Company, New York, 1959 (translated in English by K.A. Hirsch for the Russian-language book).
[15] J. van Heijenoort, “On locally convex manifolds,” Comm. Pure Appl. Math., vol. 5, no. 3, pp. 223–242, 1952.