...lysis and Control of Nonlinear Systems
...Byrnes, C.F. Martin and R.E. Saeks (editors)
...sevier Science Publishers B.V. (North-Holland), 1988

Computation of Observer Normal Form Using *Macsyma**

by Andrew R. Phelps† and Arthur J. Krener‡

Abstract

The nonlinear observer algorithm has been simplified by removal of the expensive bracket computations. Explicit solutions to the state coordinate change problem are provided. *Macsyma* was instrumental in finding these solutions and makes the algorithm readily computable.

1. Introduction

The computation of linear observers has become relatively routine, and computer packages exist which make these computations straightforward and accessible. When it comes to nonlinear observers, however, the picture has not been so bright. Algorithms for this sort of calculation have been published [1], [2], [4], [5], [6]. In general, these are limited by one or more steps involving difficult computations.

The Ph.D. thesis of Phelps [8] provides a breakthrough in the nonlinear observer algorithm. In particular, Lie bracket calculations are no longer required to perform changes of state coordinates, and the computation becomes straightforward. *Macsyma* was instrumental in developing the new approach. A prototype of this new algorithm has been implemented in *Macsyma*.

We consider an uncontrolled dynamical system with partial state observation:

$$\dot{\xi} = f(\xi) \quad \text{and} \quad y = h(\xi)\,. \tag{1}$$

The state space is in $\mathbf{R}^n$ and the output space is in $\mathbf{R}^p$. Generally, this may be put in observa*ble* normal form:

$$\dot{\xi} = A\,\xi - B\,\alpha(\xi) \quad \text{and} \quad y = C\,\xi\,. \tag{2}$$

The problem is to see if, in fact, it supports observer normal form:

$$\dot{x} = A\,x - \alpha(C\,x) \quad \text{and} \quad \bar{y} = C\,x\,. \tag{3}$$

---

* Research supported in part under AFOSR 85-0267

† University of California, Berkeley. Current address:
Dept. of Math. and Computer Sci.
San Jose State University
San Jose, CA 95192 .

‡ Address:
Dept. of Mathematics
University of California
Davis, CA 95616 .

(See [3] for detail.) Here the A, B and C are matrices given in Brunovský canonical form.

We seek to convert a system (1) to observer form (3). This paper is based on the approach in [4] and [5], as modified in [8]. The conditions (as modified) are:

Observable form Must be able to convert system to observable form (2);
Output coordinate change Must satisfy d.e. for $y = y(\bar{y})$;
Polynomial degree Observable form polynomials $f_j(\xi)$, for $1 \leq j \leq p$, (the entries of $B\,\alpha(\xi)$ in (2)) must have degree $\leq \ell_j$;
Coefficient compatibility Observable form polynomials must evaluate to certain integrals of differential expressions in injection terms (the entries of $\alpha(C\,x)$ in (3)).

Note. The coefficient compatibility condition replaces a condition that certain brackets vanish. (If q_j is the unit vector in the $\xi_{j:\ell_j}$ direction, we require that all brackets of elements in $\{ad_{-f}^{i-1} q_j \ : \ 1 \leq i \leq \ell_j\}$ must vanish.) We also take note of the approach in [1], [2] and [6], which is not used here.

2. Coefficient Compatibility in Standard Coordinates

For simplicity's sake, we first describe the results in the case that we have the "right" output $\bar{y}$. In fact, we will see that this could be considered sufficient for an improved algorithm, since the relevant change of state coordinates will be entirely determined by the tranformation $y = y(\bar{y})$ of the corresponding outputs.

In section 3, however, we will indicate a result expressed directly in observable form coordinates (ξ, y).

We may compute observ*able* form $((\bar{\xi}, \bar{y})$ coordinates), relative to the output $\bar{y}$ which is the solution to the output d.e.'s:

$$\dot{\bar{\xi}} = A\,\bar{\xi} - B\,\alpha(\bar{\xi}) \quad \text{and} \quad \bar{y} = C\,\bar{\xi}\,. \tag{4}$$

We call the coordinates (4) *standard coordinates.*

This computation does not constitute a major burden on our algorithm (given *Macsyma*), since we only require an iterated set of Lie differentiations of functions and back-substitutions to get the transformation $\xi = \xi(\bar{\xi})$.

To annotate our coordinate systems, we adopt certain conventions. We describe the state variables by $\xi_{i:j}$, indicating that it is the $(j-1)$-th time derivative of the output variable y_i, the same going for $(x, \bar{y})$ and $(\bar{\xi}, \bar{y})$ coordinates. The injection functions $\alpha_{i:j}$ are written similarly. If $p = 1$, we omit the "1 :" to simplify the notation.

Furthermore, the coefficient

$$\bar{a}_{m:\cdots}(\bar{y}) := \bar{a}_{m:\cdots|\underbrace{\cdots \underbrace{i_{j:k}+1 \cdots i_{j:k}+1}_{e_{j:k}\ \text{times}} \cdots}_{(j)}|\cdots}(\bar{y})$$

of the monomial

$$\xi^{\sharp} := \prod_{j=1}^{p} \prod_{k=1}^{r_j} \bar{\xi}_{j:i_{j:k}+1}^{e_{j:k}} \tag{5}$$

is characterized by having *degree* $i_{j:k}$ and *exponent* $e_{j:k}$ with respect to its factor $\bar{\xi}^{e_{j:k}}_{j:i_{j:k}+1}$, for $1 \le k \le r_j$ and $1 \le j \le p$. The vertical bars '|' separate the subscript into parts according to the the output j involved. We also have cumulative indices $e_j := \sum_{k=1}^{r_j} e_{j:k}$, $e := \sum_{j=1}^{p} e_j$ and $w := \sum_{j=1}^{p} \sum_{k=1}^{r_j} i_{j:k}\, e_{j:k}$. To simplify the notation, we also occasionally represent functions as their own "0-th" derivatives and we write "$\alpha_{m:i}$," for $i \le 0$, as a null symbol indicating a contribution which vanishes.

The natural way to describe the change from observable form coordinates ξ to observer form coordinates x is to compute the d.e.'s which determine it. These get quite complicated, since they involve the change of coordinates matrix $J = \dfrac{\partial y}{\partial \bar{y}}$ and its iterated derivatives. The choice of standard coordinates, however, causes all these terms to vanish.

Using, for convenience, the simplifying assumption that each observability index ℓ_j is equal to some ℓ, we compute the equations $x = x(\bar{\xi})$ governing the change of state coordinates:

$$
\begin{aligned}
\bar{\xi}_{j:1} &= x_{j:1}\,, \\
\bar{\xi}_{j:2} &= x_{j:2} - \alpha_{j:1}(\bar{y})\,, \\
&\ \ \vdots \\
\bar{\xi}_{j:\ell_j} &= x_{j:\ell_j} - \sum_{i=1}^{\ell_j-1} L_{\bar{f}}^{\ell_j-i-1}\, \alpha_{j:i}(\bar{y})\,, \\
\bar{f}_j(\bar{\xi}) &= -\sum_{i=1}^{\ell_j} L_{\bar{f}}^{\ell_j-i}\, \alpha_{j:i}(\bar{y})\,,
\end{aligned}
\tag{6}
$$

for $1 \le j \le p$, where $\bar{f}$ is f in the $\bar{\xi}$ coordinates.

From the expansion (6) we can, in effect, read off the $\bar{a}$ polynomial coefficients in terms of the α injection functions.

Example 2.1. *Coefficient solutions, for the case $p = 1$ and $\ell = 3$.*

The expansion (6) gives us:

$$f_1(\xi) = a_{23}(y)\,\xi_2\,\xi_3 + a_3(y)\,\xi_3 + a_{222}(y)\,\xi_2^3 + a_{22}(y)\,\xi_2^2 + a_2(y)\,\xi_2 + a_1(y)\,.$$

This leads to the coefficient solutions: $\bar{a}_1 = -\alpha_3$, $\bar{a}_2 = -\dfrac{d\alpha_2}{d\bar{y}}$, $\bar{a}_{22} = -\dfrac{d^2\alpha_1}{d\bar{y}^2}$, $\bar{a}_{222} = 0$, $\bar{a}_{23} = 0$ and $\bar{a}_3 = -\dfrac{d\alpha_1}{d\bar{y}}$. We integrate the $\bar{a}$'s to get the α's. △

The pattern described in the above example is easily extended to the case where we have p *equal* indices.

Theorem 2.2. *Suppose that all the indices are equal,* i.e., $\ell_j = \ell$ for $1 \le j \le p$. *Then the polynomial coefficient* $\bar{a}_{m:\cdots}(\bar{y})$ *is given by*

$$- \left(\frac{w!}{\prod_{j=1}^{p} \prod_{k=1}^{r_j} e_{j:k}!\, i_{j:k}!^{e_{j:k}}} \right) \frac{\partial^{e} \alpha_{m:\ell - w}(\bar{y})}{\partial \bar{y}_1^{e_1} \cdots \partial \bar{y}_p^{e_p}} . \tag{7}$$

The existence of an injection vector $\alpha(\bar{y})$ compatible with all the coefficients of $f_1(\xi)$, as given in (7), together with the observable form, output coordinate change and polynomial degree conditions, constitute necessary and sufficient conditions for the existence of observer normal form (supposing all indices equal).

Proof

Let m, $1 \leq m \leq p$, be fixed.

First of all, if $\ell = 1$, then (4) simply becomes $\dot{\bar{\xi}}_{m:1} = \bar{f}_m(\bar{\xi}) = \bar{f}_m(\bar{y})$. Thus $\bar{a}_{m:1} = \bar{f}_m(\bar{y}) = -\alpha_{m:1}(\bar{y})$, where we take $\alpha_{m:1} := -\bar{f}_m$. But this matches formula (7).

Note that the derivation in example 2.1 gives a result in accordance with formula (7). That illustrates the pattern we use for our induction.

As induction hypothesis, we assume that formula (7) holds for $\ell = \mu$. We target the coefficients in the expansion for $\bar{\xi}_{m:\mu} - x_{m:\mu}$ in the case $\ell = \mu + 1$ which are the source via Lie differentiation of the coefficients of $\bar{f}_m$ we seek to evaluate.

But, in the p.d.e. expansion (6), we find the same expression, $-\sum_{i=1}^{\mu} L_{\bar{f}}^{\mu - i}\, \alpha_{m:i}$, for $\bar{\xi}_{m:\mu} - x_{m:\mu}$, in the expansion with $\ell = \mu + 1$, as we find for $\bar{f}_m$ in the expansion with $\ell = \mu$. This means that the induction assumption will enable us to *know* those "target" coefficients, which, when Lie differentiated, give contributions to the expansion for $\bar{f}_m$ in the case when the multi-index ℓ is $\mu + 1$. These coefficients evaluate to nothing but the coefficients of the terms of $\bar{f}_m$ in the case (given by induction assumption) that the multi-index ℓ is μ.

Let $\xi^{\sharp}$, given as in (5), be a monomial in $\bar{f}_m(\bar{\xi})$. We wish to determine its coefficient $\bar{a}_{m:\dots}$.

An expression which under Lie differentiation by $\bar{f}$ can give a term like $\xi^{\sharp}$ may have two forms.

Case 1. It may have no increment in the exponent $e_{j:1}$ of $\bar{\xi}_{j:2}$, for $1 \leq j \leq p$.

In this instance, it will come from Lie differentiation of a term of the form

$$\xi^{:}\, \bar{\xi}_{\eta:\iota_{\eta:k}}\, \bar{\xi}^{-1}_{\eta:\iota_{\eta:k}+1} . \tag{8}$$

By induction assumption, term (8) has a numerical coefficient which calculates back from our projected coefficient for $\xi^{:}$ as

$$- \frac{\iota_{\eta:k}!\, e_{\eta:k}}{(\iota_{\eta:k} - 1)!\, e^{:}} \left(\frac{1}{w} \right) \left(\frac{w!}{\prod_{j=1}^{p} \prod_{k=1}^{r_j} e_{j:k}!\, i_{j:k}!^{e_{j:k}}} \right) .$$

Here we take $e^{\sharp} := e_{\eta:k-1} + 1$ if $\iota_{\eta:k-1} = \iota_{\eta:k} - 1$ (and $:= 1$ otherwise). The differentiation process contributes an extra factor of $e^{\sharp}$. Thus, the partial numerical contribution from this term is

$$\iota_{\eta:k}\, e_{\eta:k} \left(\frac{1}{w}\right). \tag{9}$$

In this case, the "α" part of the coefficient is carried through unchanged. Moreover, it was already of the required form, since e_η has not been altered by adding and subtracting 1 in (8).

Case 2. It may have an increment in the coefficient of $\bar{\xi}_{j:2}$ for $j = \eta$.

The source of the terms of this type is partial differentiation of the "α" coefficient by $\bar{y}_\eta$. This will give an additional factor of $\bar{\xi}_{\eta:2}$, while not affecting the numerical coefficient. The prior exponent of $\bar{\xi}_{\eta:2}$ will have been $e_{\eta:1} - 1$.

Therefore, the monomial term prior to Lie differentiation was $\xi^{\natural}\, \bar{\xi}_{\eta:2}^{-1}$. By induction assumption, its numerical contribution was

$$- \iota_{\eta:1}!\, e_{\eta:1} \left(\frac{1}{w}\right) \left(\frac{w!}{\prod_{j=1}^{p} \prod_{k=1}^{r_j} e_{j:k}!\, i_{j:k}!^{\,e_{j:k}}}\right).$$

Its partial numerical contribution is therefore

$$\iota_{\eta:1}\, e_{\eta:1} \left(\frac{1}{w}\right), \tag{10}$$

since $\iota_{j:1}! = 1 = \iota_{j:1}$ when $j = \eta$.

Note that in this case the e_η increments by 1, so that the "α" term adjusts as prescribed by formula (7).

These two cases exhaust the possibilities. The "α" coefficients are as required. And, combining (9) and (10), we get a total contribution to the numerical coefficient of a factor of

$$\left(\sum_{j=1}^{p} \sum_{k=1}^{r_j} \iota_{j:k}\, e_{j:k}\right) \left(\frac{1}{w}\right) = 1\,,$$

which is also as required. △

In the general case, where the observability indices are given arbitrarily, we adopt a recursive method for calculating the $\bar{a}$ coefficients in terms of the α injection functions.

We may take a system with arbitrary indices $\ell_1, \ldots, \ell_p$, and prolong it to a system of dimension $p \cdot \ell_p$, to which theorem 2.2 applies. Retracing the prolongation step-by-step, we can track the (increasingly complex) form of the formulas for the $\bar{a}$ coefficients. We rely on the following prolongation lemma (for a proof, see [8]):

Lemma 2.3 (Krener-Respondek-Phelps). *Suppose an uncontrolled system, given in observable form, has 2 distinct multi-indices ℓ_1, ℓ_2 of multiplicities p_1, p_2 and, further, that it may be transformed by change of output coordinates $y = y(\bar{y})$ to observer form. Then it may be prolonged to a system in observable form, having multi-indices $\lambda_1 := \ell_1 + 1$ and $\lambda_2 := \ell_2$, of the above multiplicities. Furthermore, the transformation $y = y(\bar{y})$ and the* *injection function $\alpha(\cdot)$ both prolong trivially to functions which will take the prolonged system over into observer form.*

We formulate our recursive coefficient calculation as follows:

Algorithm 2.4 (Coefficient Prolongation Algorithm). *Suppose we have solved for the coefficients belonging to the indices $\ell_1, \ldots, \ell_k$ and moreover $\ell_k < \ell_k + 1 = \ldots = \ell_s$, where $s \leq p$. We may construct a "quasi-solution" using the method of theorem 1 applied to the prolonged system with s indices all equal to ℓ_s. By back-substitution we may then express the prolonged version of $\bar{f}_j(\bar{\xi})$ with α's as coefficients, for $k+1 \leq j \leq s$. We may then substitute $L^i_{\bar{f}} \bar{\xi}_{h:\ell_h}$ (expressed in terms of the α's, and using the solutions previously derived for $1 \leq j \leq k$) for the "quasi-variables" $\bar{\xi}_{h:\ell_h+i}$ where $1 \leq i \leq \ell_s - \ell_h$. Finally, we may "read off" the coefficients of the monomials thus derived.*

We may now combine theorem 2.2 and algorithm 2.4 to get the coefficient compatibility theorem for $\bar{\xi}$ coordinates:

Theorem 2.5. *Suppose we have arbitrary indices $\ell_1, \ldots, \ell_p$. Then we may derive the coefficients $\bar{a}_{m:\ldots}(\bar{y})$ in terms of the $\alpha(\bar{y})$ injection functions by application of the Coefficient Prolongation Algorithm. The existence of an injection vector $\alpha(\bar{y})$ compatible with all the coefficients of $f_1(\xi)$, together with the observable form, output coordinate change and polynomial degree conditions, constitute necessary and sufficient conditions for the existence of observer normal form.*

Proof

This has already been done in theorem 2.2 for the case where all indices are the same. Algorithm 2.4 enables us to extend this result inductively whenever $\ell_h < \ell_{h+1}$. △

It can also be shown that we can back-solve for the α injection functions in terms of the $\bar{a}$ coefficients by iterated integrations (see [8]).

For the generic case, where there are two distinct indices, differing by 1, we state the formula:

Corollary 2.6. *For the generic case of two different multi-indices of size λ_1 and $\lambda_2 := \lambda_1 + 1$ and multiplicities p_1 and p_2, the coefficient $\bar{a}_{m:\ldots}(\bar{y})$ is given by*

$$\left(\frac{w!}{\prod_{j=1}^{p} \prod_{k=1}^{r_j} e_{j:k}!\, i_k!^{e_{j:k}}} \right) \left[\sum_{i=1}^{p_1} \left(\frac{\partial^e \alpha_{i:\lambda_1 - w}(\bar{y})}{\partial \bar{y}_1^{e_1} \cdots \partial \bar{y}_p^{e_p}} \frac{\partial \alpha_{m:\ell_m - \lambda_1}(\bar{y})}{\partial \bar{y}_i} \right) - \frac{\partial^e \alpha_{m:\lambda_2 - w}(\bar{y})}{\partial \bar{y}_1^{e_1} \cdots \partial \bar{y}_p^{e_p}} \right].$$

In particular, $\bar{a}_{m:\ldots}(\bar{y})$ is given by theorem 1 for $1 \leq m \leq p_1$ and for degree $\geq \lambda_1$ when $p_1 + 1 \leq m \leq p$.

Proof

This is directly calculated using theorem 2.2 and one application of algorithm 2.4. △

To conclude this section, we give an example of coefficients for the (simplest) non-generic case. Note that it is trivial to back-solve for the derivatives of the α's and integrate.

Example 2.7. *Some coefficient solutions, for $p = 2$, $\ell_1 = 1$, $\ell_2 = 3$.*

$$\bar{a}_{2:1} = -\alpha_{1:1}^2 \frac{\partial^2 \alpha_{2:1}}{\partial \bar{y}_1^2} - \alpha_{1:} \frac{\partial \alpha_{1:1}}{\partial \bar{y}_1} \frac{\partial \alpha_{2:1}}{\partial \bar{y}_1} + \alpha_{1:1} \frac{\partial \alpha_{2:2}}{\partial \bar{y}_1} - \alpha_{2:3},$$

$$\bar{a}_{2:|2} = 2\,\alpha_{1:1} \frac{\partial^2 \alpha_{2:1}}{\partial \bar{y}_1 \partial \bar{y}_2} + \frac{\partial \alpha_{1:1}}{\partial \bar{y}_2} \frac{\partial \alpha_{2:1}}{\partial \bar{y}_1} - \frac{\partial \alpha_{2:2}}{\partial \bar{y}_2}.$$

△

3. Coefficient Compatibility—General Case

We have seen the expansion (6) in the simplified situation of standard coordinates. This can be converted to general observable form coordinates (2) by a not-too-inconvenient calculation. However, if we have Macsyma or some other facility that enables us to do suggestive examples expeditiously, we can find patterns in the results that suggest direct solutions for the general version of (6). For instance, consider:

Example 3.1. *Coefficient solutions, in general observable form coordinates, for $p = 1$, $\ell = 4$.*

We have the following pattern, which mimics the result in theorem 2.2: $a_1 = -\frac{dy}{d\bar{y}} \alpha_4$, $a_2 = -\frac{dy}{d\bar{y}} \frac{d\alpha_3}{dy}$, $a_{23} = -3 \frac{dy}{d\bar{y}} \frac{d^2 \alpha_1}{dy^2}$, *etc.* We also have a pattern that relates to the degree ℓ terms that vanish in standard coordinates, *e.g.*:

$$a_{2222} = \frac{1}{4} \frac{d^2 a_{24}}{dy^2} - \frac{3}{16} a_{24} \frac{da_{24}}{dy} + \frac{1}{64} a_{24}^3.$$

△

To describe the intricacies of the degree ℓ terms in example 3.1, we introduce the following notational scheme. Let $P(m)$ be the partitions of m. Write a partition π of $e-1$ by $e - 1 = \sum_{j=1}^{s} c_j n_j$. Define $c := \sum_{j=1}^{s} c_j$ as the number of pieces of the partition.

With these annotations, we formulate the pattern of example 3.1 for the following theorem on coefficient compatibility:

Theorem 3.2. *A single-output system has an observer normal form iff the observable form, output coordinate change and polynomial degree conditions hold, and moreover in observable form (ξ) coordinates the coefficient $a_{\ldots}(y)$ equals*

$$-\left(\frac{w!}{\prod_{k=1}^{r} e_k!\, i_k!^{e_k}}\right) \frac{d^e \alpha_{\ell-w}(y)}{dy^e} \frac{dy}{d\bar{y}},$$

for terms of degree less than ℓ, and

$$-\sum_{\pi \in P(e-1)} \left[\left(\frac{w!\,(e-1)!}{\prod_{k=1}^{r} e_k!\, i_k!^{e_k} \prod_{j=1}^{s} c_j!\, n_j!^{c_j}} \right) \left(\frac{-1}{\ell}\right)^c \prod_{j=1}^{s} \left(\frac{d^{n_j-1} a_{2\ell}(y)}{dy^{n_j-1}} \right)^{c_j} \right],$$

for terms of degree equal to ℓ.

A proof of this theorem will appear in a forthcoming paper of Phelps [7].

Using Lemma 2.3 and the above theorem (adjusted to the case of p equal indices), we may in principle compute the general transformation $x = x(\xi)$, relating observer form (3) to observable form (2).

4. Conclusion

Two points need to be made here. First, the "coefficient compatibility" approach to nonlinear observer calculations simplifies in principle the theory and makes unwieldy bracket calculations unnecessary. Second, the use of *Macsyma* made it possible to do the rather extended calculations of examples that made the patterns in the data stand out. Every aspect of the algorithms for nonlinear observer calculation is readily accessible to *Macsyma* programming, and converting the algorithm from its abstract form of "algorithm-in-principle" to a concrete "algorithm-in-fact" is naturally done in this milieu.

References

[1] D. Bestle and M. Zeitz. Canonical form observer design for non-linear time-variable systems. *International Journal of Control*, 38(2):419–431, 1983.

[2] H. Fritz and H. Keller. Design of nonlinear observers by a two-step transformation. In M. Fliess and M. Hazewinkel, editors, *Algebraic and Geometric Methods in Nonlinear Control Theory*, D. Riedel, Dordrecht, 1986.

[3] A. J. Krener. Normal forms for linear and nonlinear systems. In M. Luksik, C. Martin, and W. Shadwick, editors, *Differential Geometry, the Interface between Pure and Applied Mathematics*, American Mathematical Society, Providence, RI. In press.

[4] A. J. Krener and A. Isidori. Linearization by output injection and nonlinear observers. *Systems & Control Letters*, 3:47–52, 1983.

[5] A. J. Krener and W. Respondek. Nonlinear observers with linearizable error dynamics. *SIAM Journal on Control and Optimization*, 23(2):197–216, 1985.

[6] C. W. Li and L. W. Tao. Observing non-linear time-variable systems through a canonical form observer. *International Journal of Control*, 44(6):1703–1713, 1986.

[7] A. R. Phelps. On computing nonlinear observers. In preparation.

[8] A. R. Phelps. *A Simplification of Nonlinear Observer Theory*. PhD thesis, University of California, Berkeley, 1987.