6/16/04 draft

Beyond Kalman Filtering: the TKF

by

Robert W. Bass

Florida Institute of Technology

ABSTRACT

In automatic control theory, Kalman's important *Separation Principle* enables the designer to first design an optimal State-Feedback Control System and, secondly, an optimal (or *unbiased minimal-variance*) State Estimator or **Kalman Filter** (**KF**), and then achieve by their combination the same as if one had used Stochastic Optimal Control Theory *ab initio*; in guidance systems this is called the *Guidance/Navigation Separation Principle* [**1**]. But in pure Navigation (or Tracking) Systems one does ***not*** always need an optimal estimate of the entire n-dimensional state-vector $x = [x_1, x_2, \cdots, x_n]'$; for example, one may need only an unbiased minimal-variance estimate of the first component x_1 of x. In such cases one may use a **Trans-Kalman Filter** (**TKF**) which obtains ***smaller*** values of the variance P_{11} of the x_1-estimate than is obtainable by *any* KF [which requires minimization of $(P_{11} + P_{22} + \cdots + P_{nn}) = trace(P)$ and of intrinsic necessity weights *identically* the variance of the estimate of *each* component of x].

Introduction

For simplicity, we shall provide details only in the case of Linear Time-Invariant (**LTI**) processes, although as usual the results may be generalized readily to include all near-to-equilibrium or near-nominal-path *non-stationary linearizeable* processes in the obvious manner, particularly if the system's *robustness margin* ρ as defined in Appendix 2 is adequately large. For further simplicity (and with an eye to the usual real-time implementation by digital computing) we suppose that the problem is considered only in Discrete Time, wherein the time-interval $dT \equiv: t_k - t_{k-1}$ between epochs $\{t_k\}$ of process-output observations, or sensor measurements, need not be constant, but is considered to be *known retroactively* at each new observation's epoch t_k , for $k = 1, 2, 3, \ldots$. Then, letting $x^k = x(t_k)$, $(k = 1, 2, 3, \ldots)$, one considers the LTI stochastic process

$$x^{k+1} = \Phi \cdot x^k + \Gamma \cdot v^k \ , \tag{1}$$

$$y^k = H \cdot x^k + w^k \ , \tag{2}$$

where Φ is a constant $n \times n$ *state-transition* matrix and Γ is a constant $n \times m$ *exogenous input-coupling* matrix or *actuator* matrix (with $m \le n$), and H is a constant $l \times n$ *output coupling* matrix or *sensor* matrix (with $l \le n$), and where the m-vector v^k and the l-vector w^k are, respectively, zero-mean Gaussian stochastic white-noise processes of given *covariance* matrices Q & R, *i.e.*

$$\boldsymbol{E}\{v^k\} \equiv 0 \ , \quad \boldsymbol{E}\{v^j \cdot (v^k)'\} = \delta_{jk} \cdot Q, \tag{3}$$

$$\boldsymbol{E}\{w^k\} \equiv 0 \ , \quad \boldsymbol{E}\{w^j \cdot (w^k)'\} = \delta_{jk} \cdot R, \tag{4}$$

where ’ denotes vector-matrix *transposition*; ***E*** denotes mathematical *expectation*; δ_{jk} denotes the Kronecker delta; and $Q = Q' \geq 0$ is a non-negative definite $m \times m$ *disturbance-covariance* matrix, while $R = R' > 0$ is a positive-definite $l \times l$ *noise-covariance* matrix. (If there is also present a known input it can be removed from consideration in (1) by the well-known Algebraic Separation Principle [*cf. e.g.* [**18**], *pp*. 62-63].)

We shall assume that the time for the Estimator Algorithm to process an output measurement y^k is negligible in comparison to dT, so that the measurement of y^{k+1} may be used virtually instantaneously to provide an online real-time *estimate* $\hat{x}^{k+1}$ of x^{k+1}. Accordingly we postulate the *information-theoretic architecture* of the KF, namely the *linear* estimator (see Rugh [**2**], Exercise 29.3, *p*. 566)

$$\hat{x}^{k+1} = \Phi \cdot \hat{x}^k + K \cdot \{ y^{k+1} - H \cdot \Phi \cdot \hat{x}^k \} \equiv$$

$$\equiv (I_n - K \cdot H) \cdot \Phi \cdot \hat{x}^k + K \cdot H \cdot \Phi \cdot x^k + K \cdot H \cdot \Gamma \cdot v^k + K \cdot w^{k+1}, \qquad (5)$$

where K is the celebrated $n \times m$ *Kalman gain matrix* and which *estimator*, when subtracted from (1) to provide the dynamics of the *estimation-error* $\tilde{x} \equiv: x - \hat{x}$, yields (after substituting new stochastic processes $\{\xi^k\}$ and $\{\eta^k\}$ by the *definitions* $\xi^k \equiv: (I_n - K \cdot H) \cdot \Gamma \cdot v^k$, $\eta^k \equiv: -K \cdot w^{k+1}$)

$$\tilde{x}^{k+1} = \Psi \cdot \tilde{x}^k + \xi^k + \eta^k, \quad \boldsymbol{E}\{\xi^k \cdot (\eta^k)'\} \equiv 0, \qquad (6)$$

where

$$\Psi \equiv: (I_n - K \cdot H) \cdot \Phi, \qquad \boldsymbol{E}\{\xi^k\} \equiv 0, \qquad \boldsymbol{E}\{\eta^k\} \equiv 0, \qquad (7a)$$

$$\boldsymbol{E}\{\xi^j \cdot (\xi^k)'\} = \delta_{jk} \cdot \tilde{Q}, \qquad \tilde{Q} \equiv: (I_n - K \cdot H) \cdot \Gamma \cdot Q \cdot \Gamma' \cdot (I_n - K \cdot H)', \qquad (7b)$$

$$\boldsymbol{E}\{\eta^j \cdot (\eta^k)'\} = \delta_{jk} \cdot \tilde{R}, \qquad \tilde{R} \equiv: K \cdot R \cdot K', \qquad (7c)$$

and where I_n denotes the $n \times n$ *identity matrix*. Note that in the continuous-time KF one must allow for the possibility of a correlation between disturbance $v(t)$ and noise $w(t)$, but in discrete time this complication happily disappears! Now define P_k, the $n \times n$ *estimation-error covariance* matrix, by $P_k \equiv: \boldsymbol{E}\{\tilde{x}^k \cdot (\tilde{x}^k)'\} \equiv (P_k)' \geq 0$, and as usual find that, for $k = 1, 2, 3, \ldots$,

$$P_{k+1} = \Psi \cdot P_k \cdot \Psi' + \tilde{Q} + \tilde{R}. \qquad (8)$$

In the LTI case it is convenient to take K to be a constant, so we postulate that P_k has a steady-state limit $P = P' \geq 0$, i.e. $P_k \to P$ as $k \to +\infty$. Thus, finally, we have

$$P = \Psi \cdot P \cdot \Psi' + \tilde{Q} + \tilde{R}. \qquad (9)$$

For convenience, convert this discrete-time *Lyapunov Equation* (9) to its continuous-time equivalent (*cf*. Grewal & Andrews [**3**], *pp*. 133-152 [assuming their $H \equiv 0$ to reduce their Riccati equation to our Lyapunov equation], or J.N. Little’s MATLAB program *dlyap* [**4**]):

$$\Psi_c \cdot P + P \cdot \Psi_c' + \tilde{Q}_c + \tilde{R}_c = 0, \qquad (10a)$$

$$\Psi_c \equiv: (\Psi + I_n)^{-1} \cdot (\Psi - I_n), \qquad (10b)$$

$$\tilde{Q}_{\mathrm{c}} \;\equiv:\; (1/2)(I_{\mathrm{n}} - \Psi_{\mathrm{c}})\cdot \tilde{Q} \cdot(I_{\mathrm{n}} - \Psi_{\mathrm{c}}'), \tag{10c}$$

$$\tilde{R}_{\mathrm{c}} \;\equiv:\; (1/2)(I_{\mathrm{n}} - \Psi_{\mathrm{c}})\cdot \tilde{R} \cdot(I_{\mathrm{n}} - \Psi_{\mathrm{c}}'). \tag{10d}$$

Let the linear operator *vec* be defined as taking an $n \times n$ matrix M and stacking its columns one above the other, in sequential order, to define an n^2-dimensional vector $\mathbf{m} \equiv: vec(M)$, and define the unique inverse linear operator *mat* such that $M \equiv: mat(\mathbf{m}) \;\equiv\; mat(vec(M))$. Also let $\otimes$ denote the *Kronecker product* (cf. Bellman [**4**], *p*. 239), so that if $I_{\mathrm{n}}\otimes\Psi_{\mathrm{c}}$ has no repeated eigenvalues then the unique solution of the Lyapunov equation (10a) is given by

$$P \;=\; vec(\mathbf{p}), \tag{11a}$$

where $\mathbf{p}$ is found by solving the non-singular system of linear equations [equivalent to (10a)]

$$[\,(\Psi_{\mathrm{c}}\otimes I_{\mathrm{n}}) \;+\; (I_{\mathrm{n}} \otimes \Psi_{\mathrm{c}}')\,]\cdot\mathbf{p} \;=\; -\,vec(\tilde{Q}_{\mathrm{c}}) \;-\, vec(\tilde{R}_{\mathrm{c}}). \tag{11b}$$

The net effect of the preceding is to define uniquely scalar functions $\tilde{\phi}$, $\tilde{\psi}$ such that

$$\tilde{\phi} \;=\; \tilde{\phi}\,(K,Q\,;\,\Phi,\Gamma,H,dT), \qquad \tilde{\psi} \;=\; \tilde{\psi}\,(K,R\,;\,\Phi,\Gamma,H,dT), \tag{12a}$$

$$P_{11} \;=\; \tilde{\phi}\,(K,Q) \;+\; \tilde{\psi}\,(K,R). \tag{12b}$$

For example, if $l \;=\; m \;=\; 1$, then

$$Q \;=\; Q_0\cdot\sigma_{\mathrm{d}}^2 \;>\; 0, \qquad R \;=\; R_0\cdot\sigma_{\mathrm{s}}^2 \;>\; 0, \tag{13a}$$

where $Q_0 \;=\; Q_0(\Phi,\Gamma,H,dT)$, $R_0 \;=\; R_0(\Phi,\Gamma,H,dT)$, and where σ_{d}^2 is the continuous-time process-disturbance spectral density and σ_{s}^2 is the continuous-time sensor-noise spectral density, and we may therefore find scalar functions ϕ and ψ such that, by repetition of the preceding arguments,

$$\phi \;=\; \phi(\,K\,;\,\Phi,\Gamma,H,dT), \qquad \psi \;=\; \psi(\,K\,;\,\Phi,\Gamma,H,dT), \tag{13b}$$

$$P_{11} \;=\; \phi(K)\cdot\sigma_{\mathrm{d}}^2 \;+\; \psi(K)\cdot\sigma_{\mathrm{s}}^2\,, \tag{13c}$$

where by inspection of (7)-(13) it may be verified that P_{11} is truly *linear* in the disturbance and sensor variances σ_{d}^2 & σ_{s}^2 as claimed.

Problem Statement

In general (at least, after a preliminary de-correlation procedure [**3**]), Q and R are diagonal in their elements $\{\,Q_{\mathrm{jj}}\,\}$ and $\{\,R_{\mathrm{kk}}\,\}$. Specifically, let $\{e^{\mathrm{j}}\}$ & $\{u^{\mathrm{k}}\}$ denote the unit column-vectors of I_{m} and I_{l}, and form the corresponding dyads $E^{\mathrm{j}} \;=\; e^{\mathrm{j}}\cdot(e^{\mathrm{j}})'$, $(j \;=\; 1, 2, \ldots , m)$ and $U^{\mathrm{k}} \;=\; = u^{\mathrm{k}}\cdot(u^{\mathrm{k}})'$, $(k \;=\; 1, 2, \ldots , l)$, so that Q is a sum of m terms $E^{\mathrm{j}}\cdot Q_{\mathrm{jj}}$ and, similarly, R is a sum of l terms $U^{\mathrm{k}}\cdot R_{\mathrm{kk}}$ where each matrix E^{j} & U^{k} has only 0's and 1's as its elements. Then a repetition of the preceding arguments, *mutatis mutandis*, establishes via linearity the existence of m scalar functions $\{\phi_{\mathrm{j}}(K) \mid (j \;=\; 1, 2, \ldots , m)\,\}$ and l scalar functions $\{\psi_{\mathrm{k}}(K) \mid (k \;=\; 1, 2, \ldots , l)\,\}$ such that

$$P_{11} = \sum_{j=1}^{m} \phi_{\mathrm{j}}(K)\cdot Q_{\mathrm{jj}} + \sum_{k=1}^{l} \psi_{\mathrm{k}}(K)\cdot R_{\mathrm{kk}} . \qquad (13\mathrm{d})$$

The TKF introduced here consists of *choosing K* to *jointly minimize* the coefficients $\{\phi_{\mathrm{j}}(K)\}$ and $\{\psi_{\mathrm{k}}(K)\}$ which display P_{11} as a linear function of the diagonal elements of Q & R. However, physically realistic values of the process-disturbance variances Q_{jj} and the sensor-noise variances R_{kk} are seldom if ever available. Accordingly it is common practice in the aerospace industry to "tune the KF" by empirical adjustment of Q & R until satisfactory performance is predicted via computer simulations (prior to actual estimator-system implementation, test, development, and deployment).

In the classical KF, not merely P_{11} but $(P_{11} + P_{22} + \cdots + P_{\mathrm{nn}}) = trace(P)$ is minimized by the choice of K, but (although this yields a stable system) it does NOT make P_{11} alone as small as it can be made when some *tolerable* degradations of the remaining P_{kk} are allowed, and there is no way short of an obvious elaboration of the theory developed herein [using (llb) and the individual parallels to (13d) for each of the remaining P_{kk}] to minimize a *weighted* sum of the P_{kk} with arbitrary weights.

Also, unfortunately, it has been found in practice that the classical KF is often quite *fragile* in the sense that $\sigma^2 = trace(P)$ as a function of $(\{Q_{\mathrm{jj}}\},\{R_{\mathrm{kk}}\})$ can have a very sharply defined minimum with severely steep walls so that when even a slightly off-nominal pair (Q,R) is encountered the resultant change in K may cause the filter's performance to degrade unacceptably from that expected. This has led to much study of possible methods to render KFs more *robust* (*e.g.* [**6**]-[**9**]), including a recent generalization to discrete time of the 1956 Bass "rhobustness" criterion ρ [**10**] presented succinctly in Appendix 2 below. But in the present context, all such efforts have hitherto suffered from the problem presented in the preceding paragraph.

Therefore we introduce here the **TKF** as an alternative approach to attempt to render the system *maximally* ***robust*** to unknown/unpredictable deviations in Q & R (which also to a considerable extent can compensate for un-modeled cross-couplings and/or discrepant deviations from the assumed design-nominal parameter-values in Φ, Γ, and H, as quantified in Appendix 2).

New Resolution

The *Trans-Kalman Filter* (**TKF**) is one in which the $(l\cdot n)$ distinct parameter-values defining the Kalman gain matrix K are so chosen as to make the $(l + m)$ nonlinear functions $(\{\phi_{\mathrm{j}}(K)\},\{\psi_{\mathrm{k}}(K)\})$ *jointly* "as small as possible" in an appropriate sense (that in each particular application may be defined on an *ad hoc* basis).

To illustrate this point in the context of the simplest possible TKF, wherein $n = 2$, and as already considered above, $l = m = 1$, consider the tracking problem ([**11**] – [**17**]) and in particular the widely used steady-state/limiting KF called an Alpha-Beta Tracker. Today's computers are so powerful and cost-effective that it is practical to use real-time nonlinear trigonometric transformations to resolve measured range, azimuth and inclination into rectilinear Euclidean coordinates, in which the dynamics of a point-particle moving in a straight line at constant velocity become uncoupled, so that two (or, in the case of a change of altitude, three) *one-dimensional* tracking problems may be considered independently. Thus, consider rectilinear constant-speed motion defined by a state-vector $x = [r,v]'$, where r denotes range and $v = dr/dt$

denotes velocity, and let r_m denote the measured range. Then the **α-β Tracker Algorithm** may be stated as

$$r_{new} = (1 - \alpha)\cdot[r_{old} + v_{old}\cdot dT] + \alpha\cdot r_m, \qquad (0 < \alpha < 2), \quad (14a)$$

$$v_{new} = (1 - \beta)\cdot v_{old} + \beta\cdot[(r_m - r_{old})/dT], \qquad (0 < \beta < 2\cdot(2 - \alpha)). \quad (14b)$$

Note that even if α & β are taken to be constants, the algorithm (14) automatically ***adapts*** to whatever the most recent update time-interval $dT \equiv: t_{new} - t_{old}$ may have been. Moreover, for fixed *dT*, (14) may be regarded as a steady-state or limiting KF by setting

$$\Phi = \begin{bmatrix} 1 & dT \\ 0 & 1 \end{bmatrix}, \qquad \Gamma\cdot Q\cdot\Gamma' = \begin{bmatrix} dT^3/3 & dT^2/2 \\ dT^2/2 & dT \end{bmatrix}\cdot\sigma_a^2, \quad (15)$$

$$H = [1, 0], \qquad K = [\alpha, \beta/dT]', \qquad R = \sigma_s^2/dT, \quad (16)$$

where σ_a^2 is the continuous-time variance of white-noise-like _acceleration disturbance_ and σ_s^2 is the continuous-time variance of the white-noise-like _sensor noise_ (as explained and derived by Kamen & Su [**17**], *pp.* 319-22). In this $n = 2$ case of (15)-(16), straightforward application of the preceding theory (1)-(11) yields, after eliminating the redundancy in the $n^2 = 4$ equations in (11), just $(1/2)\cdot n\cdot(n+1) = 3$ independent linear equations in 3 unknowns, to solve (9) for (P_{11}, P_{12}, P_{22}). After a truly enormous amount of algebra, best done by a symbolic-algebra engine, the final result simplifies greatly down to

$$P_{11} = \phi(\alpha,\beta)\cdot\sigma_a^2\cdot(dT^3/3) + \psi(\alpha,\beta)\cdot\sigma_s^2/dT, \quad (17a)$$

$$\phi \equiv: [(1 - \alpha)^2/\alpha]\cdot\{[3\cdot(2 - \alpha) - \beta]/[\beta\cdot(2\cdot(2 - \alpha) - \beta)]\}, \quad (17b)$$

$$\psi \equiv: [2\cdot\alpha^2 + (2 - 3\cdot\alpha)\cdot\beta]/[\alpha\cdot(2\cdot(2 - \alpha) - \beta)], \quad (17c)$$

Here is ψ known in the literature [*cf. e.g.* [**11**], *p.* 27, *eqn.* 1.2-20] but ϕ seems to be a novel consequence of the more complete theory (13d) introduced herein. Taking σ_s^2 to be a constant, minimizing P_{11} is the same as minimizing $\tilde{P}_{11} \equiv: P_{11}/\sigma_s^2$ so, defining the _disturbance-to-noise ratio_ ρ by

$$\rho \equiv: \sigma_a^2/\sigma_s^2, \quad (18)$$

we wish to minimize

$$\tilde{P}_{11} = \phi(\alpha,\beta)\cdot\rho\cdot(dT^3/3) + \psi(\alpha,\beta)/dT. \quad (19)$$

For each *dT* and for each “physically plausible” value of the ratio ρ , we can use standard numerical methods to find the absolute minimum of $\tilde{P}_{11}$ on the sub-triangular (α,β)-domain of assured positivity of $\tilde{P}_{11}$, namely $\{0 < \alpha < 1,\ 0 < \beta < 2\cdot(2 - \alpha)\}$, in which case it turns out that plotting α and β versus *dT* reveals that they are essentially constant for $dT \geq 1$. Accordingly one may try the “robustification” approach ([**6**]-[**10**]) by allowing ρ and *dT* to be arbitrary unknown parameters and choosing constant values of (α,β) to “jointly” minimize (ϕ,ψ) in an appropriate sense.

By inspection of (17b) it makes no sense to vary α to minimize ϕ because this immediately yields $\alpha \equiv 1$, which is tantamount to not using an α-β Tracker! [See Appendix 1 for more detailed proof.] Accordingly we regard α as fixed and focus on minimizing

$$\hat{\phi} \equiv: [3\cdot(2-\alpha)-\beta]/[\beta\cdot(2\cdot(2-\alpha)-\beta)], \tag{20}$$

Differentiating with respect to β and simplifying, one finds

$$\beta^2\cdot[2\cdot(2-\alpha)-\beta]^2\cdot(\partial\hat{\phi}/\partial\beta) = -\{\beta^2 - 6\cdot(2-\alpha)\cdot\beta + 6\cdot(2-\alpha)^2\} \equiv$$

$$\equiv -(2-\alpha)^2\cdot(\kappa^2 - 6\cdot\kappa + 6) \equiv$$

$$\equiv -(2-\alpha)^2\cdot(\kappa - 3 + 3^{1/2})\cdot(\kappa - 3 - 3^{1/2}) = 0, \tag{21}$$

wherein we were required [by inspection of the preceding quadratic form in β and the hypothesis $(\partial\hat{\phi}/\partial\beta) = 0$], to have assumed that β has the form

$$\beta = \hat{\beta}(\alpha) \equiv: \kappa\cdot(2-\alpha), \quad \kappa = 3 - 3^{1/2} = 1.267949192, \tag{22}$$

because the other possible choice of κ make β too large. Accordingly the choice (22) minimizes the consequences of large ρ or large *dT* regardless of what α-value may have been selected! Substituting (22) into

$$\hat{\psi} = \hat{\psi}(\alpha) \equiv: \psi(\alpha, \hat{\beta}(\alpha)) \tag{23}$$

yields, after differentiation and simplification,

$$(1/2)\cdot(2-\kappa)\cdot(2-\alpha)^2\cdot\alpha^2\cdot(\partial\hat{\psi}/\partial\alpha) = (2-\kappa)\cdot\alpha^2 + 4\cdot\kappa\cdot\alpha - 4\cdot\kappa = 0, \tag{24}$$

whence minimization of $\hat{\psi}$ is achieved by the choice [at first, for arbitrary κ]

$$\alpha = 2\cdot(-\kappa + [2\cdot\kappa]^{1/2})/(2-\kappa) = 0.886553879, \tag{25}$$

which shows, for *any* κ, that $\hat{\psi}(\alpha) \equiv \alpha$ [!], and then from (22), also implies that

$$\beta = 1.411793110\,. \tag{26}$$

Conclusion

The *maximally robust* (or, in complete generality, “least worst”) α-β Tracker is given by

$$\alpha_{OPT} = 0.886554, \qquad \beta_{OPT} = 1.411793\,, \tag{27}$$

$$P_{11} = (0.0243289)\cdot\sigma_a^2\cdot(dT^3/3) + \alpha_{OPT}\cdot\sigma_s^2/\mathrm{dT}, \tag{28}$$

as shown in an Numerical Appendix In this case the error-state *closed-loop* state-transition matrix becomes

$$\Psi \;=\; \begin{bmatrix} 1-\alpha & (1-\alpha)\cdot dT \\ -\beta/dT & (1-\beta) \end{bmatrix}, \tag{29}$$

whose characteristic polynomial turns out, remarkably, to be independent of *dT* [thus further justifying the choice of constant (α,β)], and is

$$z^2 - (2 - \alpha - \beta)\cdot z + (1 - \alpha) \;\equiv\; (z - z_1)\cdot(z - z_2) \;=\; 0, \tag{30}$$

which has complex roots $z_{1,2}$ ≡: $-\,0.149 \pm 0.302\cdot\boldsymbol{i}$ that satisfy [*cf.* Numerical Appendix & λ in Appendix 2]

$$|\,z_1\,| \;=\; |\,z_2\,| \;=\; 0.336818 \tag{31}$$

and therefore define poles *inside* the unit circle of the complex z-plane, ensuring <u>*tracker stability*</u> for arbitrary but known *dT* and *arbitrary/unknown* (σ_a^2, σ_s^2). Indeed, by the known results [2] cited in Appendix 2, the state-vector estimate's error-norm $\|\,\tilde{x}^{\,k}\|$ decreases by a <u>*factor*</u> of at least (1/3) [e.g. < 0.04 for $k \geq 3$] after each new update at the same time-increment *dT* !

Appendix 1

On second thought, perhaps we should have considered the possibility of taking α = 1. But in that case, since $\phi \equiv 0$ and $\psi(1,\beta) \;\equiv\; [(2\;-\beta)/(2\;-\beta)] \;\equiv\; 1$, we find that $P_{11} = \sigma_s^2/dT$, whence the value of σ_a^2 is irrelevant and β may be chosen arbitrarily, subject only to $0\; < \beta \;<\; 2$. But in the event of negligible σ_a^2 and/or small *dT* (< < 1), this result is clearly inferior to (28), so that in hindsight the decision to ignore the factor $[(1 - \alpha^2)/\alpha]$ in ϕ was correct after all.

Appendix 2

A STABILITY-RHOBUSTNESS MARGIN IN DISCRETE TIME

by

Robert W. Bass

Excerpt from a letter to Prof. Wilson J. Rugh of the Dept. of Electrical & Computer Engineering at the Johns Hopkins University dated August 13, 2002 in which I told him about my having generalized from analog time to discrete-time (in connection with an evening course on DSP for the local branch of Florida Inst. of Technology that I was then teaching) my 'Rhobustness' Criterion that I had first published for analog-time systems in 1956 [**10**] and then improved in an IEEE Conference paper [**8**] in 1991):

==

I have just derived what I believe is an elegant generalization from continuous-time systems to discrete-time systems of my theory of "rhobustification" of control and observer systems, as follows:

In your book you prove that if the system

$$x^{k+1} \;=\; A\cdot x^k + B\cdot u^k, \quad u \;=\; -\,K\cdot x,$$

is stabilized by state-feedback, to become [closed-loop]

$$x^{k+1} \;=\; A_{\mathrm{cl}}\cdot x^k, \qquad x^0 = x^{\mathrm{o}}, \qquad A_{\mathrm{cl}} \;\equiv:\; A - B\cdot K,$$

then there must be positive numbers $\gamma \;\geq\; 1$ and $\lambda \;<\; 1$ [$\lambda \equiv: \max\{|\mathrm{eig}(A_{\mathrm{cl}})|\}$] such that

$$\|x^k\| \;\leq\; \gamma\cdot\|x^{\mathrm{o}}\|\cdot\lambda^k, \qquad (k = 1,2,3,\,\ldots).$$

I now define a *stability "rhobustness" margin* ρ, $(0 < \rho < 1)$, by $\rho \equiv: (1 - \lambda)/\gamma$, and prove that if in ***NONSTATIONARY/NONLINEAR & EXTERNALLY-FORCED*** actuality the system is better-modeled by the 'perturbed' system

$$x^{k+1} = A_{cl}\cdot x^{k} + f(k, x^{k}) + g(k), \qquad (k = 1,2,3, \ldots),$$

where there exist (κ,δ) such that, for all x and k, $\|f(k, x)\| \le \kappa\cdot\|x\|$, $\|g(k)\| \le \delta$, and where $\kappa < \rho$, then

$$\|x^{k}\| \le \gamma\cdot\|x^{o}\|\cdot\Lambda^{k} + \delta/\rho \ , \quad \Lambda \equiv \{ \lambda + [\kappa/\rho](1 - \lambda)\} < 1, \quad (k = 1,2,3, \ldots).$$

Hence the ***larger*** is ρ, the ***LESS SENSITIVE*** is the actual system to *UNMODELLED* **cross-couplings** & other *NONSTATIONARY* and/or *NONLINEAR* effects, as well as *EXTERNAL FORCING!*

References

[**1**] Bucy, Richard S. & Peter D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, Chelsea, 2nd ed., 1987 [available from ***AMS*** @ **www.ams.org**].

[**2**] Rugh, Wilson J., *Linear System Theory*, 2nd ed., Prentice Hall, 1996.

[**3**] Grewal, Mohinder & Angus P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*, 2nd ed., Wiley, 2001.

[**4**] Little, J.N. in the MATLAB *Control Systems Toolbox*, The Mathworks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098.

[**5**] Bellman, Richard, *Matrix Analysis*, 2nd ed., ***SIAM***.

[**6**] Bass, Robert W., "Robustified LQG Synthesis to Specifications," *Proceedings,* Fifth Army Coordinating Group on Modern Control, Picatinny Arsenal, NJ, 1983, *pp*. B15-B93.

[**7**] Bass, Robert W. & Dean Zes, "Robust Tuning of Kalman Filters," *Proceedings,* American Control Conference (1985), Boston, MA, *pp.* 179-181.

[**8**] Bass, Robert W., "ρ-Synthesis: A 'Rhobustness' Margin for Unstructured Nonlinear and Time-Varying Deviations," *Proceedings,* 30th ***IEEE*** CDC (1991), Brighton, U.K., *pp*. 2531-2537.

[**9**] Bass, Robert W., "On 'A Tutorial Comparison of Three Multivariable Stability Margins'" *Proceedings,* First ***IEEE*** Regional Conference on Aerospace Systems, Westlake Village, CA, 1993, *pp*. 856-862.

[**10**] Bass, Robert W., "Equivalent Linearization, Nonlinear Circuit Synthesis, & the Stabilization & Optimization of Control Systems," *Proceedings*, 1956 Symposium on Nonlinear Circuit Analysis, Brooklyn Polytechnic Institute MRI Symposia Series, vol. VI (1957), *pp*. 163-198.

[**11**] Brookner, Eli, *Tracking and Kalman Filtering Made Easy*, Wiley, 1998.

[**12**] Gelb, Arthur, *ed.*, *Applied Optimal Estimation*, MIT Press, 1974.

[**13**] Chui, Charles K. & Guanrong Chen, *Kalman Filtering with Real-Time Applications*, Springer-Verlag, 1987.

[**14**] Zarchan, Paul & Howard Musoff, *Fundamentals of Kalman Filtering*, ***AIAA***, 2000.

[**15**] Bar-Shalom, Yaakov, & X. Rong Li & Thiagalingam Kirubarajan, *Estimation with Applications to Tracking and Navigation*, Wiley, 2001.

[**16**] Blackman, Samuel & Robert Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.

[**17**] Kamen, E.W. & J.K. Su, *Introduction to Optimal Estimation*, Springer-Verlag, 1999.

[**18**] Kalman, Rudolf E. & Peter L. Falb & Michael A. Arbib, *Topics in Mathematical System Theory*, McGraw-Hill, 1969.