ON A DIFFERENCE-DELAY EQUATION

Roy O. Davies
Department of Mathematics & Computer Science
University of Leicester

A.J.Ostaszewski
Department of Mathematics
London School of Economics

October 4, 1998

Abstract

We investigate how the behaviour, especially at $\pm\infty$, of continuous real solutions $f(t)$ to the equation $f(t) = a_1 f(t + h_1) + a_2 f(t - h_2)$, where a_1, a_2, h_1, h_2 are positive real constants, depends on the values of these parameters. Definitive answers are given, except in certain cases when h_1/h_2 is rational..

1. Introduction

We investigate how the behaviour of continuous real solutions $f(t)$ to the homogeneous linear difference-delay equation

$$f(t) = a_1 f(t + h_1) + a_2 f(t - h_2), \tag{$*$}$$

where

$$a_1, a_2, h_1, h_2 \text{ are positive real constants,}$$

depends on the values of these four parameters. A partial treatment of the conditions for all solutions to tend to zero as $t \to +\infty$ was supplied in [1] in connection with a problem in economic theory. Now we shall deal with a wider range of types of behaviour, such as are frequently in question for solutions to differential and difference equations, giving complete answers except in certain cases when h_1/h_2 is rational. (We propose to go more fully into these in a subsequent paper [2].)

Specifically, we investigate the conditions under which some or all solutions of $(*)$ are small (that is, tend to zero) as $t \to \pm\infty$; some or all (apart from the trivial solution $f(t) \equiv 0$) are unbounded as $t \to \pm\infty$; some solution is non-oscillatory (that is, ultimately non-zero) as $t \to \pm\infty$; and in this last case, whether there is one that tends to zero, or one that tends to $+\infty$.[There always exist non-trivial oscillatory solutions, for example of the form $f(t) = e^{at} \cos bt$.]

Of course any linear combination of solutions is also a solution, including $f(t) \equiv 0$, and moreover any function $f(t)$ continuous on an interval $T - h_2 \leq t \leq T + h_1$ extends to a solution of $(*)$ on $\mathbf{R}$, provided that $(*)$ holds at $t = T$; we may, and shall, treat all solutions as defined on the whole of $\mathbf{R}$. However, for most types of behaviour we clearly cannot expect conclusions that apply to all solutions or the whole of $\mathbf{R}$.

As we shall see, among the solutions of $(*)$ there will be some of the form $f(t) = e^{zt}$ (or the real or imaginary part, if z is complex) for all values of the parameters, and additionally some of the form $f(t) = te^{zt}$ for particular values; the nature of such solutions depends on whether z is real or complex and upon the sign of $Re(z)$, and turns out to be an excellent guide to what can be asserted of solutions in general. However, in the absence of any useful representation of the general solution in terms of these special exponential solutions, we have been forced to use (§§ 2-6)direct real-variable methods; those we have developed share some common themes and do seem to be of interest in themselves.

We now record the effect of a few simple transformations. If f satisfies $(*)$ and $g(t) = f(-t)$ then g satisfies the same equation except that (a_1, h_1) and (a_2, h_2) are interchanged; therefore results about behaviour as $t \to -\infty$ can be read off from those about behaviour as $t \to +\infty$, and we may concentrate on the latter. On the other hand, if $\lambda > 0$ and $g(t) = f(\lambda t)$ then g satisfies $(*)$ with h_1, h_2 replaced by $H_1 = h_1/\lambda, H_2 = h_2/\lambda$ respectively; therefore where convenient we may suppose that $H_1 + H_2 = 1$ (by taking $\lambda = h_1 + h_2$), whilst any case with h_1/h_2 rational can be reduced to one in which h_1, h_2 are integers. Finally, if we set $f(t) = e^{\mu t} g(t)$ where $\mu \neq 0$ then g satisfies $(*)$ with a_1, a_2 replaced by $a_1 e^{\mu h_1}, a_2 e^{-\mu h_2}$; we shall make use of this transformation too, and of course it could in addition be used to obtain, from our results, answers to certain questions about the order of magnitude of solutions.

Unless otherwise stated, solutions of $(*)$ will always be understood as real and continuous on the whole of $\mathbf{R}$.

For the convenience of anyone seeking answers to specific questions, we now list the principal topics of the remaining sections.

§2. Existence of a non-oscillatory solution
§3. Existence of a divergent solution
§4. Existence of a small non-oscillatory solution
§5. Boundedness or smallness of all solutions: sufficient conditions
§6. Non-existence of bounded or small non-trivial solutions
§7. Solutions e^{zt} when h_1/h_2 is irrational
§8. Smallness of all solutions: general remarks

2. Existence of a non-oscillatory solution

We establish the following conclusive result.

THEOREM 1. *Equation* $(*)$ *admits a solution that is non-oscillatory as* $t \to +\infty$ *or* $t \to -\infty$ *if and only if*

$$a_1^{h_2} a_2^{h_1} \leq h_1^{h_1} h_2^{h_2} / (h_1 + h_2)^{h_1 + h_2}. \qquad (**)$$

Moreover, when this condition is satisfied there exists a solution of the form $f(t) = e^{xt}$, *x real, and thus positive for all* $t \in \mathbf{R}$; *when* $(**)$ *holds with equality* *there is a unique such solution $f(t) = e^{x_0 t}$, but $f(t) = te^{x_0 t}$ is also a solution (and hence $f(t) = (At + B)e^{x_0 t}$ for any constants A, B).*

2.1. *Proof of sufficiency*

For $f(t) = e^{xt}$ to be a solution it is necessary and sufficient that $\varphi(x) = 1$, where

$$\varphi(x) = a_1 e^{h_1 x} + a_2 e^{-h_2 x}. \tag{2.1}$$

This function is strictly convex on $\mathbf{R}$ and has a unique minimum at $x = X$, where X is given by

$$a_1 h_1 e^{h_1 X} = a_2 h_2 e^{-h_2 X}, \tag{2.2}$$

and therefore, $\varphi(x) = 1$ has a real solution if and only if $\varphi(X) \leq 1$, which reduces to $(**)$. If $(**)$ holds with inequality then $\varphi(X) < 1$ and the equation $\varphi(x) = 1$ will have two solutions, while if $(**)$ holds with equality then $\varphi(X) = 1$ as well as $\varphi'(X) = 0$, and the equation $\varphi(x) = 1$ will have the unique solution $x = X$. In the latter case we have $a_1 e^{h_1 X} + a_2 e^{-h_2 X} = 1$ as well as (2.2) and it follows that $f(t) = te^{Xt}$ is a solution of $(*)$ in addition to $f(t) = e^{Xt}$.

2.2. *Reformulation of necessity*

Suppose that $(**)$ is false, and let $\varphi(x)$ be as in (2.1), X be given by (2.2), and $\lambda = \varphi(X)$; then $\lambda > 1$, and it is easy to verify that if f satisfies $(*)$ and we set $f(t) = e^{Xt} f^*(t)$ then

$$f^*(t) = \lambda \left[\frac{h_2}{h_1 + h_2} f^*(t + h_1) + \frac{h_1}{h_1 + h_2} f^*(t - h_2) \right].$$

From this concavity condition we must deduce that f^* cannot be positive for (say) all large positive t. We establish the following somewhat stronger result.

THEOREM 2. *There exists no function $f(t)$ which for all large t is defined, positive, and continuous, and satisfies*

$$f(t) \geq \lambda \left[\frac{h_2}{h_1 + h_2} f(t + h_1) + \frac{h_1}{h_1 + h_2} f(t - h_2) \right] \tag{2.3}$$

where λ, h_1, h_2 are positive constants with $\lambda > 1$.

2.3. *Proof of Theorem 2*

Suppose that f is such a function, which we shall call *anomalous.* We may and shall suppose that $h_1 + h_2 = 1$ (see §1). Now for all large A and $t \geq A$ we have $f(t) > 0$ and

$$\int_A^t f(\tau)d\tau \geq \lambda \left[h_2 \int_{A+h_1}^{t+h_1} f(\tau)d\tau + h_1 \int_{A-h_2}^{t-h_2} f(\tau)d\tau \right], \tag{2.4}$$

that is, with $F(t)$ denoting the left side of (2.4),

$$F(t) \geq \lambda[h_2 F(t+h_1) + h_1 F(t-h_2)] + \lambda \left[h_1 \int_{A-h_2}^{A} f(\tau)d\tau - h_2 \int_A^{A+h_1} f(\tau)d\tau \right]. \tag{2.5}$$

We observe first that if f is strictly decreasing (as well as anomalous), then

$$\lambda \left[h_1 \int_{A-h_2}^{A} f(\tau)d\tau - h_2 \int_A^{A+h_1} f(\tau)d\tau \right] > \lambda[h_1 h_2 f(A) - h_2 h_1 f(A)] = 0,$$

and (2.5) shows that F itself is an anomalous function, as well as being (ultimately) strictly increasing. We now consider two cases.

CASE 1. $\displaystyle\int_A^\infty f(\tau)d\tau < +\infty$. Then integration of (2.3) shows that

$$\int_t^\infty f(\tau)d\tau \geq \lambda \left[h_2 \int_{t+h_1}^\infty f(\tau)d\tau + h_1 \int_{t-h_2}^\infty f(\tau)d\tau \right],$$

and thus $G(t) = \displaystyle\int_t^\infty f(\tau)d\tau$ is an anomalous function, as well as being strictly decreasing. As observed above, there then exists a strictly increasing anomalous function.

CASE 2. $\displaystyle\int_A^\infty f(\tau)d\tau = +\infty$. Since $h_1 + h_2 = 1$, we can rewrite (2.5) as

$$F(t) - c \geq \lambda\{h_2[F(t+h_1) - c] + h_1[F(t-h_2) - c]\}$$

where

$$c = c(A) = -\frac{\lambda}{\lambda - 1} \left[h_1 \int_{A-h_2}^{A} f(\tau)d\tau - h_2 \int_A^{A+h_1} f(\tau)d\tau \right].$$

Since $F(t)$ is strictly increasing and $F(t) \to +\infty$ as $t \to +\infty$, we conclude that $F(t) - c$ is a strictly increasing anomalous function.

Thus in both cases there exists a strictly increasing anomalous function; let $g(t)$ denote one such. Evidently $\int_t^\infty g(\tau)d\tau = +\infty$, and the discussion for Case 2 above shows that for sufficiently large A_1, A_2 there exist constants c_1, c_2 such that the function

$$G(t) = \int_{A_1}^{t} \left[\int_{A_2}^{\tau_1} g(\tau_2)d\tau_2 - c_2 \right] d\tau_1 - c_1$$

is anomalous; moreover G is twice differentiable, with $G''(t) = g(t) > 0$ for all large t.

The inequality (2.3), applied to G, gives

$$0 > (1-\lambda)G(t) \geq \lambda(h_2[G(t+h_1) - G(t)] + h_1[G(t-h_2) - G(t)]).$$

By the mean value theorem, the right side can be written as

$$\lambda h_1 h_2 [G'(t_2) - G'(t_1)]$$

where $t - h_2 < t_1 < t < t_2 < t + h_1$, and this again can be written as

$$\lambda h_1 h_2 (t_2 - t_1) G''(t_3)$$

where $t_1 < t_3 < t_2$. But $G''(t_3) = g(t_3) > 0$ (if t is sufficiently large), so we have a contradiction. ■

REMARK. The above proof evidently shows more generally that if $(**)$ is false then the inequality

$$f(t) \geq a_1 f(t+h_1) + a_2 f(t-h_2)$$

admits no solution for large positive t which is non-oscillatory and is assumed locally Lebesgue integrable rather than continuous We may return to the question of the behaviour of measurable solutions of $(*)$, if there turns out to be anything essentially new to say.

3. Existence of a divergent solution

A solution of $(*)$ will be called *divergent* at $\pm\infty$ if $|f(t)| \to +\infty$ as $t \to \pm\infty$. By continuity, every such solution is non-oscillatory; it was shown in §2 that a non-oscillatory solution exists if and only if $(**)$ holds, and that when this condition is satisfied there exists a solution of the form $f(t) = e^{xt}$ where x is real. However this will be divergent at $+\infty$ (resp. $-\infty$) only if $x > 0$ (resp. $x < 0$). We shall establish a necessary and sufficient condition for the existence of a divergent solution.

THEOREM 3 *Equation $(*)$ admits a solution divergent at $+\infty$*
(i) *when $a_1 + a_2 < 1$,*
(ii) *when $a_1 + a_2 = 1$ if and only if $a_1h_1 \le a_2h_2$,*
(iii) *when $a_1 + a_2 > 1$ if and only if $a_1h_1 < a_2h_2$ and $(**)$ holds.*
Moreover, if one of these conditions is satisfied then there exists a solution of the form $f(t) = e^{xt}, x > 0$, except when both $a_1 + a_2 = 1$ and $a_1h_1 = a_2h_2$, in which case every linear function $f(t) = At + B$ is a solution.

As indicated in §1, the corresponding result at $-\infty$ can be deduced immediately by interchanging (a_1, h_1) and (a_2, h_2); thus it simply involves reversing the inequalities between a_1h_1 and a_2h_2 in (ii) and (iii). We pass to the proof of Theorem 3.

3.1. *Proof of sufficiency*

Consider the graph of the function $\varphi(x)$ defined by equation (2.1), for which $\varphi(0) = a_1 + a_2$, $\varphi'(0) = a_1h_1 - a_2h_2$, and $(**)$ is equivalent to $\min\{\varphi(x) : x \in \mathbf{R}\} \le 1$. It is easy to see that (i), (ii), (iii) give the conditions for the equation $\varphi(x) = 1$ to admit a positive root, and therefore $(*)$ to admit a solution $f(t) = e^{xt}$ with $x > 0$, with the stated exception $a_1 + a_2 = 1$ and $a_1h_1 = a_2h_2$, which is precisely when the linear functions are solutions.

3.2. *Proof of necessity*

This can be reduced, *via* a careful consideration of cases (which it seems unnecessary to set forth), and taking Theorem 1 into account, to showing that *assuming* $a_1 + a_2 \geq 1$ *and* $a_1 h_1 > a_2 h_2$, *there is no solution divergent at* $+\infty$.

We shall use an appropriate adaptation of the method of §2. Suppose if possible that there does exist a solution divergent at $+\infty$, which we may and shall assume to be ultimately positive. As in §2, we conclude that there are constants A_1, A_2, c_1, c_2 such that with

$$F(t) = \int_{A_1}^{t} \left[\int_{A_2}^{\tau_1} f(\tau_2) d\tau_2 - c_2 \right] d\tau_1 - c_1$$

we have

$$F(t) = a_1 F(t + h_1) + a_2 F(t - h_2)$$

and F is strictly increasing and positive, with $F''(t) = f(t) > 0$. But because $a_1 h_1 > a_2 h_2$, for large t we have

$$F(t) > a_1 F(t + h_1) + a_2 F\left(t - \frac{a_1 h_1}{a_2}\right)$$

and so because $a_1 + a_2 \geq 1$ (and by the mean value theorem)

$$\begin{aligned}
0 &\geq (1 - a_1 - a_2) F(t) \\
&\geq a_1 \left[F(t + h_1) - F(t) \right] - a_2 \left[F(t) - F\left(t - \frac{a_1 h_1}{a_2}\right) \right] \\
&= a_1 h_1 \left[F'(t_2) - F'(t_1) \right] \quad \text{where} \ \ t - \frac{a_1 h_1}{a_2} < t_1 < t < t_2 < t + h_1 \\
&= a_1 h_1 (t_2 - t_1) F''(t_3) \quad \text{where} \ \ t_1 < t_3 < t_2 \\
&> 0, \quad \text{a contradiction.} \ \blacksquare
\end{aligned}$$

4. Existence of a small non-oscillatory solution

A solution of $(*)$ will be called *small* at $\pm\infty$ if $f(t) \to 0$ as $t \to \pm\infty$. Conditions under which there exists at least one non-trivial small solution will be discussed later; here we establish necessary and sufficient conditions for the existence of a non-oscillatory small solution.

THEOREM 4 *Equation* $(*)$ *admits a solution small and non-oscillatory at* $+\infty$

(i) *when* $a_1 + a_2 < 1$,

(ii) *when* $a_1 + a_2 = 1$, *if and only if* $a_1 h_1 \geq a_2 h_2$,

(iii) *when* $a_1 + a_2 > 1$, *if and only if* $a_1 h_1 > a_2 h_2$ *and* $(**)$ *holds.*

Moreover, if one of these conditions is satisfied then there exists a solution of the form $f(t) = e^{xt}, x < 0$.

As usual, the result at $-\infty$ can be obtained by reversing the inequality between $a_1 h_1$ and $a_2 h_2$. We pass to the proof of Theorem 4.

4.1. *Proof of sufficiency*

A consideration of the graph of the function $\varphi(x)$ defined by (2.1) would now show that (i), (ii), (iii) are precisely the conditions under which $(*)$ admits a solution $f(t) = e^{xt}$ with $x < 0$. Unlike in Theorem 3, equality is not permitted in (ii): this is because although linear functions can be divergent, they cannot be small and non-oscillatory.

4.2. *Proof of necessity*

This can be reduced to showing that

if either (i) $a_1 + a_2 = 1$ *and* $a_1 h_1 \leq a_2 h_2$, *or* (ii) $a_1 + a_2 > 1$ *and* $a_1 h_1 < a_2 h_2$, *then there exists no solution which is both small and non-oscillatory at* $+\infty$.

Since distinct variants of the method of §2 appear to be needed, we shall treat cases (i) and (ii) separately. From the existence of a solution f which is small and non-oscillatory at $+\infty$, which we assume to be ultimately positive, a contradiction will be derived.

4.3. *Case* (i): $a_1 + a_2 = 1$ *and* $a_1 h_1 \le a_2 h_2$.

Since $f(t) > 0$ and $f(t) \to 0$ as $t \to +\infty$, the integral $\int_{t_1}^{t_2} f(\tau)d\tau$ tends to zero as $t_1, t_2 \to +\infty$ with $|t_1 - t_2| \le C$, and is positive for $t_1 < t_2$. We may suppose that $f(t) > 0$ for $t \ge -h_2$. Consider the expression

$$G(t) = a_1 \int_t^{t+h_1} \left(\int_0^{\tau_2} f(\tau_1)d\tau_1 \right) d\tau_2 - a_2 \int_{t-h_2}^{t} \left(\int_0^{\tau_2} f(\tau_1)d\tau_1 \right) d\tau_2 .$$

By the mean value theorem we can write

$$G(t) = (a_1 h_1) \int_0^{t_2} f(\tau)d\tau - (a_2 h_2) \int_0^{t_1} f(\tau)d\tau \quad \text{where} \quad t - h_2 < t_1 < t < t_2 < t + h_1$$

$$= (a_1 h_1) \int_{t_1}^{t_2} f(\tau)d\tau - (a_2 h_2 - a_1 h_1) \int_0^{t_1} f(\tau)d\tau. \tag{4.1}$$

On the other hand

$$\begin{aligned} G'(t) &= \left[\int_0^{t+h_1} f(\tau)d\tau - \int_0^{t} f(\tau)d\tau \right] - a_2 \left[\int_0^{t} f(\tau)d\tau - \int_0^{t-h_2} f(\tau)d\tau \right] \\ &= a_1 \int_t^{t+h_1} f(\tau)d\tau - a_2 \int_{t-h_2}^{t} f(\tau)d\tau , \end{aligned}$$

so that

$$G'(t) \to 0 \ as \ t \to +\infty. \tag{4.2}$$

Also

$$G''(t) = a_1 f(t + h_1) - (a_1 + a_2) f(t) + a_2 f(t - h_2) = 0,$$

since $(*)$ holds and $a_1 + a_2 = 1$. Therefore $G'(t)$ is equal to a constant, which by (4.2) is zero, so $G(t)$ is itself constant.

If $a_1 h_1 = a_2 h_2$, this contradicts (4.1), since the right side is positive but tends to zero as $t \to +\infty$.

If $a_1 h_1 < a_2 h_2$, then by letting $t_1 \to +\infty$ in (4.1) we deduce that

$$G(t) = -(a_2 h_2 - a_1 h_1) \int_0^{+\infty} f(\tau)d\tau,$$

a finite number, and (4.1) can be written

$$-(a_2h_2 - a_1h_1)\int_{t_1}^{+\infty} f(\tau)d\tau = (a_1h_1)\int_{t_1}^{t_2} f(\tau)d\tau,$$

which is impossible because the left side is negative but the right side positive.

REMARK 1. The argument here may appear a little more direct than in §2 and §3, inasmuch as it involves, besides f , only the single auxiliary function G, but it is closely related, since $G(t) = a_1F(t+h_1) - (a_1+a_2)F(t) + a_2F(t-h_2)$ where $F(t) = \int_0^t \left(\int_0^{\tau_2} f(\tau_1)d\tau_1\right) d\tau_2$.

REMARK 2. In the case $a_1 + a_2 = 1$, $a_1h_1 = a_2h_2$ it is indeed possible for equation $(*)$ to admit *oscillatory* small solutions. For example, if $h_1 = 1,\ h_2 = 2, a_1 = 2/3, a_2 = 1/3$ then $f(t) = 2^{-t}\cos \pi t$ is a solution.

4.4. *Case* (ii): $a_1 + a_2 > 1$ *and* $a_1h_1 < a_2h_2$.

As before, we may suppose that $f(t) > 0$ for $t \geq -h_2$. Consider first the expression

$$\begin{aligned} F_1(t) &= a_1\int_0^{t+h_1} f(\tau)d\tau - \int_0^t f(\tau)d\tau + a_2\int_0^{t-h_2} f(\tau)d\tau \\ &= a_1\int_t^{t+h_1} f(\tau)d\tau - a_2\int_{t-h_2}^t f(\tau)d\tau + (a_1+a_2-1)\int_0^t f(\tau)d\tau. \end{aligned} \tag{4.3}$$

Since $F_1'(t) = a_1f(t+h_1) - f(t) + a_2f(t-h_2) = 0$, therefore $F_1(t)$ is a constant; by passing to the limit in (4.3) we conclude that

$$F_1(t) = (a_1 + a_2 - 1)\int_0^{+\infty} f(\tau)d\tau,$$

a finite number, and (4.3) can be written

$$(a_1 + a_2 - 1)\int_t^{+\infty} f(\tau)d\tau = a_1\int_t^{t+h_1} f(\tau)d\tau - a_2\int_{t-h_2}^t f(\tau)d\tau. \tag{4.4}$$

Now let $f_1(t) = \int_t^{+\infty} f(\tau)d\tau$. Then (4.4) (or integration of $(*)$) shows that f_1 satisfies the same equation $(*)$ as f; moreover f_1 is strictly decreasing and tends to zero as $t \to +\infty$. Therefore the preceding derivation of (4.4) applies equally well to f_1. With the help of the mean value theorem, the corresponding equation yields

$$(a_1+a_2-1)\int_t^\infty f_1(\tau)d\tau = (a_1h_1)f_1(t_2)-(a_2h_2)f_1(t_1) \text{ where } t-h_2 < t_1 < t < t_2 < t+h_1,$$

which is a contradiction, the left side being positive but the right-side negative, since f_1 is strictly decreasing and $a_1h_1 < a_2h_2$. ∎

5. Boundedness or smallness of all solutions: sufficient conditions

We shall establish the following result.

THEOREM 5. *If $a_1 \geq a_2+1$ then every solution of $(*)$ is bounded on $\mathbf{R}^+$, and indeed every solution is small at $+\infty$ unless both*

$$a_1 = a_2 + 1 \tag{5.1}$$

and

$$h_1/h_2 = p/q \text{ where } p \text{ is an even and } q \text{ an odd integer.} \tag{5.2}$$

[If (5.1) and (5.2) hold and K is the common value of h_1/p and h_2/q, then $f(t) = \sin(\pi t/K)$ is a non-small solution of $(*)$.]

As will be proved in §7, if h_1/h_2 is irrational then the condition $a_1 \geq a_2+1$ is also necessary, for every solution to be bounded on $\mathbf{R}^+$; the situation when h_1/h_2 is rational will be dealt with in our proposed second paper [2].

Results for $t \to -\infty$ can be deduced in the usual way; see also the Corollary at the end of §6.

5.1. *Boundedness*

Let f be a continuous real solution of $(*)$ and let

$$\sup\{|f(t)| : 0 \leq t \leq h_1 + h_2\} = M;$$

then a simple double induction shows that

$$|f(t)| \leq \left(\frac{a_2+1}{a_1}\right)^n M \quad for \quad t \geq n(h_1+h_2),\ for\ n = 1, 2, \ldots.$$

Thus f is small at $+\infty$ whenever $a_1 > a_2 + 1$. *Henceforth we assume* (5.1). We now have $|f(t)| \leq M$ for $t \geq 0$, and boundedness is established.

5.2. *Local oscillation*

For $\delta > 0$, let

$$w(\delta) = \sup\{|f(t+\eta) - f(t)| : 0 \leq \eta \leq \delta \ and\ 0 \leq t \leq h_1 + h_2\}; \tag{5.3}$$

then $w(\delta) \to 0$ as $\delta \to 0+$, and since each function $f_\eta(t) = f(t+\eta) - f(t)$ satisfies the same equation $(*)$ as $f(t)$, it follows from the reasoning in §5.1 that

$$|f(t+\eta) - f(t)| \leq w(\delta)\ for\ 0 \leq \eta \leq \delta\ and\ t \geq 0. \tag{5.4}$$

5.3. *Completion of the proof*

Suppose if possible that (5.2) is false but f is a continuous real solution of $(*)$ that is not small at $+\infty$. Then as $t \to +\infty$

$$\limsup |f(t)| = A > 0; \tag{5.5}$$

we may assume that

$$\limsup f(t) = A, \tag{5.6}$$

and we shall derive a contradiction. Let

$$b = \max\left(a_1 + a_2, \frac{a_1+1}{a_2}\right); \tag{5.7}$$

then $b > 1$ by (5.1). Choose $\delta > 0$ so small that $w(\delta) < A$; then by §5.2

$$t_1, t_2 \geq 0 \text{ and } |t_1 - t_2| \leq \delta \text{ imply } \ |f(t_1) - f(t_2)| < A. \tag{5.8}$$

The first key idea is that if t is large and the value of $f(t)$ is close to A (resp. $-A$) then so are the values of $f(t-h_1)$ and $-f(t-h_1-h_2)$. We now present this in a more precise form.

LEMMA. *Suppose that* $0 < b\eta < 1$ *and*

$$f(T) \geq (1-\eta)A, \ |f(T-h_1)| \leq (1+\eta)A, \ |f(T-h_1-h_2)| \leq (1+\eta)A. \tag{5.9}$$

Then

$$f(T-h_1) \geq (1-b\eta)A \ \text{ and } \ -f(T-h_1-h_2) \geq (1-b\eta)A. \tag{5.10}$$

If (5.9) holds with $-f$ *for* f *then so does (5.10).*

Proof From $(*)$ for $t = T - h_1$ we have (using (5.1), (5.7), and (5.9))

$$\begin{aligned} f(T-h_1) &= a_1 f(T) + a_2 f(T-h_1-h_2) \\ &\geq a_1(1-\eta)A - a_2(1+\eta)A \\ &= [1-(a_1+a_2)\eta]A \\ &\geq (1-b\eta)A, \end{aligned}$$

and likewise

$$\begin{aligned} -f(T-h_1-h_2) &= \frac{1}{a_2}[a_1 f(T) - f(T-h_1)] \\ &\geq \frac{a_1}{a_2}(1-\eta)A - \frac{1}{a_2}(1+\eta)A \\ &= \left[1 - \left(\frac{a_1+1}{a_2}\right)\eta\right] A \geq (1-b\eta)A. \end{aligned}$$

Replacing f by $-f$ establishes the final assertion.

The second key idea is that if $\epsilon > 0$ is sufficiently small, T sufficiently large, and $f(T) \geq (1-\epsilon)A$, then repeated application of the lemma will take us back to points of the form $T - mh_1$ and $T - n(h_1 + h_2)$ with n odd, at which the values of f are not too far from A and $-A$ respectively, but provided (5.2) is false these two points can be brought so close together that (5.8) is contradicted. We now pass to the details.

Because (5.2) is false, we can find positive integers m, n with $m \geq n$ and n odd such that

$$|mh_1 - n(h_1 + h_2)| < \delta. \tag{5.11}$$

(The left side may be zero.) Choose $\epsilon > 0$ so small that

$$b^m \epsilon < \frac{1}{2}. \tag{5.12}$$

By (5.5) and (5.6) we can choose $t_0 \geq m(h_1 + h_2)$ so large that both

$$t \geq t_0 - m(h_1 + h_2) \textit{ implies } |f(t)| \leq (1+\epsilon)A$$

and

$$f(t_0) \geq (1-\epsilon)A.$$

Noting that $b > 1$, so that $|f(t)| \geq (1+\epsilon)A$ implies $|f(t)| \leq (1 + b^r\epsilon)A$ for $r \geq 1$, we conclude by repeated applications of the lemma, and using (5.12), that

$$f(t_0 - mh_1) \geq (1 - b^m\epsilon)A \geq \frac{1}{2}A,$$

$$-f(t_0 - n(h_1 + h_2)) \geq (1 - b^n\epsilon)A \geq \frac{1}{2}A,$$

which on addition gives a contradiction to (5.8), in view of (5.11). ■

6. Non-existence of bounded or small non-trivial solutions

It was shown in §5 that if $a_1 \geq a_2 + 1$ then every (continuous real) solution of $(*)$ is bounded on $\mathbf{R}^+$, and indeed small at $+\infty$ except when $a_1 = a_2 + 1$ and h_1/h_2 takes certain rational values (even/odd). It will be shown in §7 that the situation changes if $a_1 < a_2 + 1$, in that unbounded solutions will now exist except in special cases; however, bounded and indeed small non-trivial solutions will also continue to exist if a_1 is not too much less than $a_2 + 1$ (see §8). We now show that this ceases to be true when a_1 drops below the value $a_2 - 1$.

THEOREM 6 *If $a_1 \leq a_2 - 1$ then no non-trivial solution of $(*)$ is small at $+\infty$, and indeed every non-trivial solution is unbounded on $\mathbf{R}^+$ unless both*

$$a_1 = a_2 - 1 \tag{6.1}$$

and

$$h_1/h_2 = p/q \quad \text{where } p \text{ is an odd and } q \text{ an even integer.} \tag{6.2}$$

[*If (6.1) and (6.2) hold and K is the common value of h_1/p and h_2/q, then $f(t) = \sin(\pi t/K)$ is a bounded non-trivial solution of $(*)$.*]

As in the case of Theorem 5, we refer to §7 for a proof that if h_1/h_2 is irrational then the condition $a_1 \leq a_2 - 1$ is also necessary, for the non-existence of non-trivial small solutions, and to §8 for a discussion of the rational case. There are corresponding results for $t \to -\infty$.

The proof of Theorem 6 is based on much the same idea as that of Theorem 5, although some of the details are simpler; we shall set it out in a similar way.

6.1. *Absolute value considerations*

If f is a continuous real solution of $(*)$ then

$$|f(t-h_2)| = \frac{1}{a_2}|f(t) - a_1 f(t+h_1)| \leq \left(\frac{a_1+1}{a_2}\right) \max\{|f(t)|, |f(t+h_1)|\},$$

and a simple double induction shows that for all t

$$|f(t)| \leq \left(\frac{a_1+1}{a_2}\right)^n \limsup_{t\to+\infty}|f(t)|, \quad \text{for } n = 1, 2, \dots\ .$$

Thus if $a_1 \le a_2 - 1$ and f is small at $+\infty$ (i.e. lim sup $|f(t)| = 0$) or $a_1 < a_2 - 1$ and f is bounded on $\mathbf{R}^+$ (i.e. lim sup $|f(t)| < +\infty$) then f must be identically zero; this proves part of Theorem 6. *Henceforth we assume that (6.1) is true and (6.2) is false*; to complete the proof we must derive a contradiction from the existence of a non-trivial solution which is bounded on $\mathbf{R}^+$. Thus we assume that $0 < \limsup_{t\to+\infty}|f(t)| = A < +\infty$, and it now follows that

$$|f(t)| \le A \text{ for all } t; \tag{6.3}$$

we may also assume that

$$\limsup_{t\to+\infty} f(t) = A. \tag{6.4}$$

6.2. *Lipschitz condition*

The function $f_1(t) = \int_t^{t+c} f(\tau)d\tau$, where $c > 0$, satisfies the same equation $(*)$ as f and is bounded, and since f is not identically zero neither is f_1, provided that c is (for example) sufficiently small. Moreover f_1 satisfies the Lipschitz condition $|f_1(t_1) - f_1(t_2)| \le 2A|t_1 - t_2|$. Thus we may assume that for some constant $L > 0$

$$|f(t_1) - f(t_2)| \le L.|t_1 - t_2| \text{ for all } t_1, t_2. \tag{6.5}$$

The idea behind the proof of Theorem 6 is the same as for Theorem 5, except that we now move forwards from a "bad" value t_0 of t instead of backwards, and obtain a conflict with the more precise (6.5) in place of (5.4).

6.3. *Persistence of values near* $\pm A$

LEMMA *If $0 < \eta < 1$ then $f(T) > (1-\eta)A$ implies*

$$f(T + h_2) > \left(1 - \frac{a_2}{a_1}\eta\right) A \text{ and } -f(T + h_1 + h_2) > \left(1 - \frac{a_2}{a_1}\eta\right) A.$$

The same implication holds with f replaced by -f.

Proof If $f(T) > (1-\eta)A$ then from $(*)$ for $t = T + h_2$, by (6.1) and (6.3)

$$\begin{aligned} f(T+h_2) &= a_1 f(T+h_1+h_2) + a_2 f(T) \\ &> -a_1 A + a_2(1-\eta)A \\ &= (1-a_2\eta)A \geq \left(1 - \frac{a_2}{a_1}\eta\right) A, \end{aligned}$$

and likewise

$$\begin{aligned} -f(T+h_1+h_2) &= \frac{1}{a_1}\left[a_2 f(T) - f(T+h_2)\right] \\ &> \frac{a_2}{a_1}(1-\eta)A - \frac{1}{a_1}A \\ &= \left(1 - \frac{a_2}{a_1}\eta\right) A. \end{aligned}$$

This remains valid if f is replaced by $-f$.

6.4. *Completion of the proof*

Because (6.2) is false, there exist positive integers m, n with $m \geq n$ and n odd such that

$$|mh_2 - n(h_1+h_2)| < A/L. \tag{6.6}$$

Choose $\varepsilon > 0$ so small that

$$\left(\frac{a_2}{a_1}\right)^m \varepsilon < \frac{1}{2}. \tag{6.7}$$

By (6.4) we can choose a value t_0 such that $f(t_0) > (1-\varepsilon)A$. By repeated applications of the Lemma we conclude using (6.7) that

$$f(t_0 + mh_2) > \left[1 - \left(\frac{a_2}{a_1}\right)^m \varepsilon\right] A > \frac{1}{2}A,$$

$$-f(t_0 + n(h_1+h_2)) > \left[1 - \left(\frac{a_2}{a_1}\right)^n \varepsilon\right] A > \frac{1}{2}A,$$

which on addition gives a contradiction to (6.5), in view of (6.6).

6.5. *Consequence of Theorems 5 and 6*

From these theorems and corresponding results at $-\infty$, among other things the following conclusions can be drawn, showing (in conjunction with Theorem 7 in the next section) that what can be said about certain types of behaviour for all solutions of the form $Re(e^{zt})$ is a faithful guide to what is true for general solutions.

COROLLARY *Every non-trivial solution of* $(*)$ *is small at* $+\infty$ *and bounded at* $-\infty$ *if* $a_1 > a_2 + 1$, *and vice versa if* $a_2 > a_1 + 1$; *when* h_1/h_2 *is irrational, the assertions remain true if* $a_1 = a_2 + 1$ *or* $a_2 = a_1 + 1$ *respectively.*

7. Solutions of the form e^{zt} when h_1/h_2 is irrational

The next theorem supplies *inter alia* a converse to the Corollary at the end of §6.

THEOREM 7 (i) *If* $a_1 < a_2 + 1$ *and* h_1/h_2 *is irrational then* $(*)$ *admits a complex solution of the form* $f(t) = e^{zt}$ *where* $Re(z) > 0$.

(ii) *If* $a_2 < a_1 + 1$ *and* h_1/h_2 *is irrational then* $(*)$ *admits a complex solution of the form* $f(t) = e^{zt}$ *where* $Re(z) < 0$.

We note first that the two assertions are equivalent, because if we set $f^*(t) = f(-t)$, then f satisfies $(*)$ if and only if f^* satisfies

$$f^*(t) = a_2 f^*(t + h_2) + a_1 f^*(t - h_1).$$

Thus it will be sufficient to prove (i).

Clearly $f(t) = e^{zt}$ is a solution of $(*)$ if and only if

$$1 = a_1 e^{h_1 z} + a_2 e^{-h_2 z}. \tag{7.1}$$

One method of proof, assuming Bloch's Theorem from complex function theory, was described in [1]. Here we shall adopt an alternative 'real-variable' approach, closely related to one method for attacking the case of rational h_1/h_2.

7.1. *Solving a simpler problem*

If $z = x + iy$ $(x, y$ real) then (7.1) becomes

$$1 = a_1 e^{h_1 x} e^{i h_1 y} + a_2 e^{-h_2 x} e^{-i h_2 y}, \tag{7.2}$$

and if $h_1 y = \beta, h_2 y = \alpha$ then this can be written

$$1 = a_1 e^{h_1 x} e^{i\beta} + a_2 e^{-h_2 x} e^{-i\alpha}. \tag{7.3}$$

We shall show that if $a_1 < a_2 + 1$ then there is a non-empty open interval $I \subseteq \mathbf{R}^+$ such that for $x \in I$, (7.3) can be solved for α and β, regarded as independent, and then that for a dense set of values of $x \in I$ we can choose α, β to satisfy $\alpha/h_2 = \beta/h_1 = y$, say: now $z = x + iy$ will be the required solution of (7.1).

First we present a sufficient condition for solvability of the equation

$$1 = Ae^{i\beta} + Be^{-i\alpha}. \tag{7.4}$$

LEMMA 1 *If* $A > 0, B > 0$, *and there exists a triangle with sides 1,A,B, then there exist numbers* $\alpha, \beta \in (0, \pi)$ *satisfying (7.4), given by*

$$\cos\alpha = (B^2 + 1 - A^2)/2B \quad \textit{and} \quad \cos\beta = (A^2 + 1 - B^2)/2A;$$

moreover

$$\sin\alpha = H/B \quad \textit{and} \quad \sin\beta = H/A,$$

where H>0 is given by

$$4H^2 = (A + B - 1)(A + 1 - B)(B + 1 - A)(A + B + 1).$$

Proof In the triangle, let α, β be the angles opposite A and B respectively.

7.2. *Solvability of (7.3)*

Henceforth we assume that $a_1 < a_2 + 1$, and for $x \geq 0$ we set

$$A(x) = a_1 e^{h_1 x}, \quad B(x) = a_2 e^{-h_2 x};$$

in view of the preceding discussion, we shall need the following result.

LEMMA 2 *There exists a non-empty open interval $I \subseteq \mathbf{R}^+$ such that $x \in I$ implies that $1, A(x), B(x)$ are the sides of a triangle.*

Proof We require that

$$A(x) < B(x) + 1, \quad 1 < A(x) + B(x), \quad B(x) < A(x) + 1. \tag{7.5}$$

By hypothesis, the first condition holds at $x = 0$, and since $A(x) - B(x) \to +\infty$ as $x \to +\infty$ there exists $x_1 > 0$ such that it holds for $0 < x < x_1$, while $A(x_1) = B(x_1) + 1$. The other two conditions are automatically satisfied at x_1, and therefore all three are satisfied throughout an interval $I = (x_0, x_1)$ where $0 < x_0 < x_1$.

7.3. *Completion of the proof*

In view of the lemmas, there exists a non-empty open interval $I \subseteq \mathbf{R}^+$ such that for all $x \in I$ (7.5) holds and there exist numbers $\alpha(x), \beta(x) \in (0, \pi)$ satisfying the conditions of Lemma 1 (with variables replaced by corresponding functions), and in particular

$$1 = a_1 e^{h_1 x} e^{i\beta(x)} + a_2 e^{-h_2 x} e^{-i\alpha(x)}. \tag{7.6}$$

Now tedious but straightforward calculations show that

$$h_2\beta'(x) - h_1\alpha'(x) = -\{h_1^2 A(x)^2 + h_1 h_2[B(x)^2 + A(x)^2 - 1] + h_2^2 B(x)^2\}/H(x),$$

and (from (7.5)) that the quadratic form in h_1, h_2 is positive definite for $x \in I$. Therefore the continuous function $h_2\beta(x) - h_1\alpha(x)$ is strictly decreasing in I. Since h_1/h_2 is irrational, numbers of the form $(h_1 l - h_2 k)\ 2\pi$, for $k,\ l \in \mathbf{Z}$, are dense in $\mathbf{R}$. Therefore for a dense set of values of $x \in I$ we have

$$h_2\beta(x) - h_1\alpha(x) = (h_1 l - h_2 k)2\pi \ \textit{where}\ k, l \in \mathbf{Z}. \tag{7.7}$$

But then with

$$y = \frac{1}{h_1}[\beta(x) + 2k\pi] = \frac{1}{h_2}[\alpha(x) + 2l\pi]$$

it follows from (7.6) and (7.7) that (7.2) is satisfied, as required. ■

8. Smallness of all solutions: general remarks

We finally return to the question that originally motivated much of the present investigation: what are the necessary and sufficient conditions on a_1, a_2, h_1, h_2 (each assumed positive) for all solutions of (*) to be small at $+\infty$? [Dual results concerning unboundedness of all non-trivial solutions or behaviour at $-\infty$ instead of $+\infty$ can of course be found in a similar way or deduced easily.] It follows from the results of the preceding sections that *if h_1/h_2 is irrational then the condition $a_1 \geq a_2 + 1$ is necessary and sufficient*.

The situation when h_1/h_2 is rational is more complicated. In the exceptional case (5.2) of Theorem 5, strict inequality is necessary as well as sufficient, because if $h_1/h_2 = p_1/p_2$ where p_1 is an even integer and p_2 an odd integer, and K is the common value of h_1/p_1 and h_2/p_2, and and if $a_1 \leq a_2 + 1$ then there exists $x \geq 0$ with $a_1 e^{h_1 x} = a_2 e^{h_2 x} + 1$, and $f(t) = e^{xt} \sin(\pi t/K)$ is a non-small solution of $(*)$.

Now suppose that h_1/h_2 is rational but not of the 'exceptional' form; by passing to a function of the form $f(\lambda t)$ where $\lambda > 0$ (compare Section 1) we can reduce our problem to one in which h_1, h_2 *are positive integers with no common factor and h_1 is odd*, so this can be assumed. We wish to determine the region A in the open first quadrant of the (a_1, a_2)-plane for which all solutions of $(*)$ are small at $+\infty$. By Theorems 3,5, and 6, this contains none of the points at which $a_1 + a_2 < 1$ or $a_2 > a_1 + 1$ (the shaded regions in Fig. 1) but all of the triangular region T in which $a_1 \geq a_2 + 1$; in our proposed second paper [2] we shall show that A extends beyond the boundary line $a_1 = a_2 + 1$ into part of the central strip S, and we shall describe methods whereby A can be specified for any particular values of h_1 and h_2. In particular, if $h_1 = h_2 = 1$ then A includes the half of S lying below the line $a_2 = a_1$, while if $h_1 = 1, h_2 = 2$ then A includes the part of S lying below the curve $a_2^2 + a_2 = a_1^2$.

Place FIGURE 1 here

References

[1] S. Gomulka, A. Ostaszewski, R.O. Davies, The innovation rate and Kalecki's theory of trend, unemployment and the business cycle, *Economica* vol. 57 (1990),525-540.

[2] Roy O. Davies and A.J. Ostaszewski, Small solutions to a difference-delay equation in the rational case, in preparation.