Canad. Math. Bull. Vol. **15** (4), 1972

ON STURM'S SEPARATION THEOREM

BY
PAUL R. BEESACK

1. **Introduction.** The purpose of this note is to obtain an extension of the classical Sturm separation theorem for the second order, linear selfadjoint differential equation

(1) $$(ry')'+sy = 0,$$

to the case of a noncompact interval. The classical theorem (cf. [3, p. 209], [4, p. 224]) assumes that r and s are continuous with r positive on a compact interval I, and concludes that between each pair of zeros (on I) of one (nontrivial) solution of (1) there lies precisely one zero of any other linearly independent solution of (1). If I is not compact, a function y which is a solution of (1) on I may often be extended (continuously) to an endpoint, say a, of I. In case $y(a)=0$, does the separation theorem hold? A complete answer to this question is given in Theorem 1 below, and an application of the extended theorem is given in Theorem 2.

2. **The separation theorem for noncompact intervals.** We shall assume that r and s are continuous, and r is (strictly) positive on an open interval (a, b), where $-\infty \leq a < b \leq +\infty$. We say that a is *singular* for the equation (1) if either $a=-\infty$, or if a is finite but either of $\lim_{x\to a+} r(x)$ or $\lim_{x\to a+} s(x)$ do not exist in the real numbers, or if $\lim_{x\to a+} r(x)=0$. Otherwise, a is said to be *nonsingular*. In other words, a is nonsingular if and only if a is finite and both r and s can be extended to continuous functions on $[a, b)$ with r positive on this interval. Similarly, b is nonsingular for (1) if and only if b is finite and r, s can be extended in the same way to $(a, b]$. Any point of (a, b) is also called nonsingular for (1).

In the case that a or b is nonsingular for (1), solutions y of (1) can also be extended so that y and y' are continuous at a or b respectively, and satisfy (1) on the extended interval. This is a consequence of the standard existence theorem for (1). For completeness we shall formulate our theorem to include the classical case; this is part (b) of the theorem and we shall subsequently refer to this part as "the separation theorem".

THEOREM 1. *Let y_1, y_2 be two linearly independent solutions of* (1) *on* (a, b). *Then we have* (*Abel's identity*)

(2) $$r(x)[y_1(x)y_2'(x)-y_2(x)y_1'(x)] \equiv k \neq 0 \quad \textit{for} \quad a < x < b,$$

Received by the editors November 3, 1970 and, in revised form, January 20, 1971.

https://doi.org/10.4153/CMB-1972-086-7 Published online by Cambridge University Press

where k is a constant. Suppose that x_1 and $\bar{x}_1$, where $a \le x_1 < \bar{x}_1 \le b$, are consecutive zeros of y_1. Then

(a) *y_2/y_1 is strictly monotonic on $(x_1, \bar{x}_1)$.*

(b) *If x_1, $\bar{x}_1$ are both nonsingular for (1), y_2 has precisely one zero on $(x_1, \bar{x}_1)$.*

(c) *If x_1 is nonsingular, $\bar{x}_1 = b$ is singular, and if $y_2(x) \ne 0$ for $x_1 < x < b$, then we have $y_2(x) = 0[y_1(x)]$ as $x \to b-$, so that $y_2(b) = 0$. Moreover, this case can occur, as can the cases that y_2 has a (single) zero on (x_1, b) with $y_2(b) = 0$ or with $y_2(b) \ne 0$.*

(d) *If $\bar{x}_1$ is nonsingular and $x_1 = a$ is singular, and if $y_2(x) \ne 0$ for $a < x < \bar{x}_1$, then we have $y_2(x) = 0[y_1(x)]$ as $x \to a+$, so that $y_2(a) = 0$. This case also occurs as do the cases that y_2 has a (single) zero on $(a, \bar{x}_1)$ with $y_2(a) = 0$ or $y_2(a) \ne 0$.*

(e) *If $x_1 = a$ and $\bar{x}_1 = b$ are both singular for (1), and if $y_2(x) \ne 0$ for $a < x < b$, then either $y_2(x) = 0[y_1(x)]$ as $x \to a+$, or $y_2(x) = 0[y_1(x)]$ as $x \to b-$, or both. All three of these cases can occur, as well as cases in which y_2 has a (single) zero on (a, b) and either no, one, or two zeros at the singular endpoints.*

Proof. We have, for all $x \in (a, b)$,

$$(r(x)y_2'(x))' + s(x)y_2(x) \equiv 0,$$
$$(r(x)y_1'(x))' + s(x)y_1(x) \equiv 0.$$

On multiplying the first of these identities by $y_1(x)$, the second by $y_2(x)$, and subtracting, we obtain

$$y_1(x)(r(x)y_2'(x))' - y_2(x)(r(x)y_1'(x))' \equiv 0,$$

or

$$\frac{\mathrm{d}}{\mathrm{d}x}\{y_1(x)(r(x)y_2'(x)) - y_2(x)(r(x)y_1'(x))\} \equiv 0,$$

from which Abel's identity (2) follows. Moreover, if $k = 0$, then since $r(x) \ne 0$ on (a, b) it would follow that $y_1y_2' - y_2y_1' \equiv 0$, or $(d/dx)(y_2/y_1) \equiv 0$, so $y_2 \equiv Ay_1$ at least on $(x_1, \bar{x}_1)$. This contradicts the hypothesis that y_1, y_2 are linearly independent on (a, b). Thus $k \ne 0$ in (2). We note that if a is nonsingular, then r, y_1, and y_2 can be extended to a so that (2) also holds for $x = a$. Similarly if b is nonsingular we may assume that (2) holds for $x = b$. Finally, we note that if y_1, y_2 are linearly dependent on (a, b) so $y_2 \equiv Ay_1$, then $k = 0$ in (2) as can be seen by direct substitution.

To prove (a), we note that (between consecutive zeros of y_1) we can write (2) in the form

$$\frac{\mathrm{d}}{\mathrm{d}x}\left\{\frac{y_2(x)}{y_1(x)}\right\} = \frac{k}{r(x)y_1^2(x)}, \qquad x_1 < x < \bar{x}_1.$$

Since $r(x)y_1^2(x) > 0$ on $(x_1, \bar{x}_1)$ while $k \ne 0$, it follows that y_2/y_1 is strictly monotonic on $(x_1, \bar{x}_1)$. Similarly of course, y_1/y_2 is strictly monotonic between consecutive zeros of y_2.

Suppose now that x_1, $\bar{x}_1$ are both nonsingular for (1). Then we may assume that y_1, y_2' are both continuous at x_1 and $\bar{x}_1$ and, in addition, that $y_2(x_1) \ne 0$, $y_2(\bar{x}_1) \ne 0$.

For, if $y_2(x_1)=0(=y_1(x_1))$ for example, then setting $x=x_1$ in (2) would give $k=0$; similarly for $\bar{x}_1$. If now we had $y_2(x)\neq 0$ for $x_1<x<\bar{x}_1$, hence also for $x_1\leq x\leq \bar{x}_1$, then by part (a), y_1/y_2 would be strictly monotonic on $[x_1, \bar{x}_1]$. This is impossible since y_1/y_2 is zero at both endpoints of this interval. Hence $y_2(x_2)=0$ for at least one point $x_2\in(x_1, \bar{x}_1)$. Interchanging the roles of y_1, y_2 we see that y_2 can have no other zeros on $(x_1, \bar{x}_1)$ proving the separation theorem (b).

The proofs of (c) and (d) are essentially the same, so we prove (c). Suppose that $y_2(x)\neq 0$ for $x_1<x<b$, hence for $x_1\leq x<b$. With no loss of generality we may assume that both y_1 and y_2 are positive on (x_1, b). By (a), y_1/y_2 is then a positive function which is strictly increasing on (x_1, b) since it is zero at x_1. This clearly implies that $y_1(x)>Ky_2(x)$ for a positive constant K and $x\geq x_1'>x_1$, proving the main part of (c). We postpone the examples for the last parts of (c), (d) and (e) to the end.

To prove the main part of (e), note that if y_2 has no zeros on (a, b), then by (a), y_1/y_2 is either strictly increasing or strictly decreasing on (a, b). Since $y_1(a)=y_1(b)=0$, it follows as in (c) and (d) that either $y_2(x)=0[y_1(x)]$ as $x\to b-$ or $y_2(x)=0[y_1(x)]$ as $x\to a+$.

The proof of the theorem will be complete when we construct the (twelve) examples illustrating the variety of cases in parts (c)–(e). For (c) and (d) the two differential equations (with the general solutions indicated)

$$((1+x^2)^2y')'+2(1+x^2)y=0, \qquad y=(A+Bx)(1+x^2)^{-1}, \tag{3}$$

$$((1+x^2)y')'+4(1+x^2)^{-1}y=0, \qquad y=A\sin(2\operatorname{Arctan} x+B), \tag{4}$$

are sufficient to give us our examples. For (c), we take $x_1=0$ and $b=\bar{x}_1=\infty$. Then, using equation (3), $y_1=x(1+x^2)^{-1}$ has consecutive zeros at the nonsingular point x_1 and the singular point $\bar{x}_1$, but $y_2=(1+x^2)^{-1}$ has no zeros on $(0, \infty)$ while $y_2(\infty)=0$. On the other hand, using the same y_1 but taking $y_2=(1-x)(1+x^2)^{-1}$, we see that y_2 has a single zero (at $x=1$) on $(0, \infty)$, with $y_2(\infty)=0$. For the remaining case of (c), we use equation (4) with $y_1=\sin(2\operatorname{Arctan} x)$ which also has consecutive zeros at $x_1=0$, $\bar{x}_1=\infty$, while $y_2=\sin(2\operatorname{Arctan} x-\pi/2)$ has a single zero (at $x=1$) on $(0, \infty)$ with $y_2(\infty)=1$. Corresponding examples for (d) are obtained by replacing the interval $(0, \infty)$ by $(-\infty, 0)$ in the preceding discussion.

For the six examples required by (e), we examine the four differential equations

$$((1+x)^2y')'+\tfrac{1}{4}x^{-2}y=0, \qquad y=\{x/(x+1)\}^{1/2}\{A+B\ln(x/(x+1))\}, \tag{5}$$

$$((1+x^2)^{1+\alpha}y')'+2\alpha(1+x^2)^\alpha y=0, \qquad y=(1+x^2)^{-\alpha}\left\{A+B\int_{-\infty}^{x}(1+t^2)^{\alpha-1}\,dt\right\}, \tag{6}$$

$$((1+x^2)y')'+(1+x^2)^{-1}y=0, \qquad y=A\sin(\operatorname{Arctan} x+B), \tag{7}$$

$$(x^{1/2}(1+x^2)^2y')'+3x^{1/2}(1+x^2)y=0, \qquad y=(1+x^2)^{-1}(A+Bx^{1/2}). \tag{8}$$

In equations (5) and (8), we use $(a, b)=(0, \infty)$, while in (6) and (7) we take $(a, b)=(-\infty, \infty)$. (Note that $x=0$ is singular in (8) only because $r(0)=0$.) First, take $y_1=\{x/(x+1)\}^{1/2}\ln\{x/(x+1)\}$ in (5), so that y_1 has consecutive zeros at 0 and ∞; the solution $y_2=\{x/(x+1)\}^{1/2}$ of (5) has no zeros on $(0,\infty]$ but has $y_2(0)=0$. By making the trivial change of variable $x=-t$ in (5), we obtain an example with $(a,b)=(-\infty,0)$ in which $y_1=\{t/(t-1)\}^{1/2}\ln\{t/(t-1)\}$ has consecutive zeros at $-\infty$ and 0, while $y_2=\{t/(t-1)\}^{1/2}$ has no zeros on $[-\infty,0)$ but has $y_2(0)=0$. In equation (6) we assume that $0<\alpha<\frac{1}{2}$ (so the integral exists) and note that $y_1=(1+x^2)^{-\alpha}$ is a solution having consecutive zeros at $-\infty$ and ∞, while $y_2=(1+x^2)^{-\alpha}\int_{-\infty}^{x}(1+t^2)^{\alpha-1}\,dt$ has no zeros on $(-\infty,\infty)$ but has $y_2(\pm\infty)=0$. (In this case we note that *all* solutions of (6) have zeros at $\pm\infty$.) For equation (7), the solution $y_1=\cos(\operatorname{Arctan} x)$ has consecutive zeros at $-\infty$ and ∞ while $y_2=\sin(\operatorname{Arctan} x)$ has a single zero on $(-\infty,\infty)$ at $x=0$, and $y_2(\pm\infty)=\pm 1$. Equation (8) has the solution $y_1=x^{1/2}(1+x^2)^{-1}$ with consecutive zeros at 0 and ∞, while the solution $y_2=(1-x^{1/2})(1+x^2)^{-1}$ has a single zero on $(0,\infty)$ at $x=1$ and $y_2(0)=1$, $y_2(\infty)=0$. For our final example, we return to equation (6) and note that for all $\alpha>0$, $y_1=(1+x^2)^{-\alpha}$ has consecutive zeros at $-\infty$, ∞ while the solution $y_2=(1+x^2)^{-\alpha}\int_0^x(1+t^2)^{\alpha-1}\,dt$ has a single zero on $(-\infty,\infty)$ at $x=0$, and has zeros at both singular endpoints.

3. **An application.** The following theorem generalizes to noncompact intervals Lemma 1.4 of [1]. This lemma was used in [1] to obtain generalizations of Wirtinger's inequality. In a paper to appear [2], similar results are obtained on noncompact intervals using Theorem 2. This theorem deals with a solution y_1 of (1) which has *three* consecutive zeros a, $\bar{x}$, and b on an interval. The main point of the theorem is to replace the extreme pair of zeros a, b and the solution y_1 by an "equivalent" pair of zeros and a solution Y such that Y has no *other* zeros on (a,b), whereas y_1 has an (interior) zero at $\bar{x}$.

THEOREM 2. *Let r and s be continuous with r positive on (a,b) where $-\infty\le a<b\le\infty$, and suppose the differential equation* (1) *has a solution y_1 with consecutive zeros at a, $\bar{x}$, b where $a<\bar{x}<b$. Let u be a function which is continuous on $[a,b]$ with $u(a)=u(b)$. Then either every solution of* (1) *which vanishes at a point of (a,b) has only a single zero on (a,b) and zeros at both a and b, or there exists a solution Y of* (1) *and two points x_1, x_2 with $a\le x_1<\bar{x}\le x_2<b$ (or $a<x_1\le\bar{x}<x_2\le b$) such that $u(x_1)=u(x_2)$, and $Y(x_1)=Y(x_2)=0$, while $Y(x)\ne 0$ for any other points of (a,b).*

Proof. First we note that if $a=-\infty$, the hypothesis on u means that $\lim_{x\to a+}u(x)=u(a)$ exists (finite), and equals $u(b)$. Now, for each $t\in(a,\bar{x}]$ denote by $y_t(x)$ the unique solution of (1) satisfying the initial conditions $y_t(t)=0$, $y_t'(t)=1$. Thus, for instance, $y_{\bar{x}}(x)=c_1y_1(x)$ for some constant $c_1\ne 0$. By parts (b), (c) of Theorem 1, if $a<t<\bar{x}$, then $y_t(x)$ vanishes precisely once on $(\bar{x},b]$ and not at b if b is nonsingular, and vanishes either at b alone or at a single point of $(\bar{x},b)$

and possibly also at b if b is singular. In either case, *we let* $T(t)$ *denote the first zero of* y_t *to the right of* t, *so that* $\bar{x}<T(t)\leq b$ *always holds*. By parts (b) and (d) of Theorem 1, $y_t(x)$ *has no zeros on* $(a, \bar{x}]$ *except at* $x=t$, while $y_t(a)\neq 0$ if a is non-singular, and $y_t(a)$ may be zero if a is singular.

Either (i) $T(t)=b$ for $a<t\leq\bar{x}$, or (ii) there exists $t_1\in(a, \bar{x})$ such that $\bar{x}<T(t)<b$ for $a<t<t_1$ and $T(t)=b$ for $t_1<t\leq\bar{x}$, or (iii) $\bar{x}<T(t)<b$ for $a<t<\bar{x}$, but $T(\bar{x})=b$ of course. To see that case (ii) is the only alternative to (i) and (iii), it suffices to note that if $T(t_0)=b$ holds for some $t_0\in(a, \bar{x})$, then b is necessarily singular for (1), and—by the separation theorem—if $t\in(t_0, \bar{x})$ then $y_t(x)\neq 0$ for $x\in(t, b)$; but then $y_t(b)=0$ since y_t must have a zero on $(t, b]$ as noted above. Thus, if $T(t_0)=b$, then $T(t)=b$ must hold for $t_0\leq t\leq\bar{x}$ which establishes the alternative (ii). Later we shall show that $T(t_1)=b$ must hold, so that case (iii) can be subsumed into case (ii), where now $t_1\in(a, \bar{x}]$. (Note that $t_1=\bar{x}$ always occurs when b is non-singular, and may occur when b is singular.)

We observe that (i) means that b is singular, and that every solution y of (1) (being a linear combination of y_1 and y_{t_0} $(a<t_0<\bar{x})$) has a zero at b. Moreover, if y vanishes at a point $t\in(a,\bar{x}]$, then $y=cy_t$ for some $c\neq 0$, so y has no other zeros on (a, b). Let y be a nontrivial solution of (1) which has a zero at a point $\tilde{x}\in(\bar{x}, b)$; such solutions *do* exist. Then y has no zeros on $(\tilde{x}, b)$ or on $(\bar{x}, \tilde{x})$ by the separation theorem. Since y can have no zeros on $(a, \bar{x}]$ by what was just proved (otherwise y would be a multiple of some y_t), it follows that y has precisely one zero on (a, b). Moreover, by part (d) of Theorem 1, $y(a)=0$. But then every solution of (1) is a linear combination of y_1 and y and so has a zero at a (as well as one at b). This completes the proof of the first alternative in the conclusion of the theorem.

Turning now to the (enlarged) case (ii), we note that the function T is continuous on (a, t_1); moreover, by the separation theorem T is strictly increasing on (a, t_1). It follows that $\lim_{t\to a+} T(t)=\bar{x}_1$ exists, with $\bar{x}\leq\bar{x}_1<b$, and that $\lim_{t\to t_1-} T(t)=b$. To prove the second assertion, we note that the strictly increasing character of T on (a, t_1), together with $\bar{x}<T(t)<b$, implies that $\lim_{t\to t_1-} T(t)=b_1$ exists, and $\bar{x}<b_1\leq b$. Suppose $b_1<b$. Choose any $x_2\in(b_1, b)$ and let y be a nontrivial solution of (1) such that $y(x_2)=0$. By the separation theorem (with y_1), y has no other zeros on $[\bar{x}, b)$. However, for each $t_0\in(a, t_1)$, y must have precisely one zero, say t_α, on $(t_0, T(t_0))$ since $T(t_0)<b_1$. This is impossible since if $t_1<t_\alpha(<\bar{x})$, then $x_2=T(t_\alpha)=b$ which is not the case, while if $t_\alpha\leq t_1$ then either $x_2=T(t_\alpha)<b_1$ or $x_2=b$ neither of which is the case. It follows that $b_1=b$. Similarly, it now follows that $T(t_1)=b$ as asserted earlier. For if not, then $T(t_1)<b$ and hence $\lim_{t\to t_1-} T(t)\leq T(t_1)<b$, a contradiction.

In case $\bar{x}_1>\bar{x}$, *we note that for any nontrivial solution* Y *of* (1) *which vanishes at a point* $x_2\in(\bar{x}, \bar{x}_1]$, *we have* $Y(a)=0$ *but* $Y(x)\neq 0$ *on* (a, b) *except at* x_2. To prove this we note first that $Y(x)\neq 0$ on (x_2, b) or on $[\bar{x}, x_2)$ by the separation theorem. Moreover, $Y(x)\neq 0$ on $(a, \bar{x})$ since if we had $Y(t_0)=0$ for some $t_0\in(a, \bar{x})$, then the solutions $y_t(x)$ for $a<t<t_0$ would have consecutive zeros at t and at $T(t)>$

$\bar{x}_1 \geq x_2$, contradicting the separation theorem. Finally, since a is necessarily singular when $\bar{x}_1 > \bar{x}$, it follows from part (d) of Theorem 1 that $Y(a)=0$, completing the proof of the assertion. We note also in this case that *every* solution of (1) has a zero at a.

Now, let u be any function which is continuous on $[a, b]$ with $u(b)=u(a)$. Suppose first that $u(\bar{x})=u(a)$. In this case the second alternative of the theorem follows on setting $x_1=a$, $x_2=\bar{x}$ and $Y(x)\equiv y_1(x)$. (In case $u(x)>u(a)$ on $(a,\bar{x})$ and $u(x)<u(a)$ on $(\bar{x}, b)$, or vice-versa, it is clear that $Y=ky_1$ is the only possible choice.) If $u(\bar{x})\neq u(a)$, then without loss of generality *we may assume that* $u(\bar{x})<u(a)$. If, in addition, $u(a)\leq u(\bar{x}_1)$, it follows from the continuity of u that there is at least one point $x_2 \in (\bar{x}, \bar{x}_1]$ such that $u(x_2)=u(a)$. On taking $x_1=a$ and $Y(x)$ to be any nontrivial solution of (1) which vanishes at x_2, it follows from the last paragraph that $Y(x_1)=Y(x_2)=0$ and $Y(x)\neq 0$ for any other points of (a, b), while $u(x_1)=u(x_2)$ by the choice of x_2.

To complete the proof of the theorem, it remains to consider the case $u(\bar{x}_1)<u(a)$ (and $u(\bar{x})<u(a)$). To handle this case, we define the function h by

$$h(t) = u[T(t)]-u(t), \qquad a < t \leq \bar{x}.$$

h is continuous on (a, t_1), and $\lim_{t\to t_1-} T(t)=b=T(s)$ for $t_1\leq s\leq \bar{x}$, while $\lim_{t\to a+} T(t)=\bar{x}_1$. Hence, if we define $h(a)=u(\bar{x}_1)-u(a)$, then h becomes continuous on $[a, \bar{x}]$ with $h(a)<0$, and $h(\bar{x})=u(b)-u(\bar{x})=u(a)-u(\bar{x})>0$. It follows that h has at least one zero on $(a, \bar{x})$, say $h(x_1)=0$. We now set $x_2=T(x_1)$ and $Y(x)\equiv y_{x_1}(x)$. Then $0=h(x_1)=u(x_2)-u(x_1)$, while $Y(x_1)=Y(x_2)=0$ and Y has no other zeros on (a, b). Since $a<x_1<\bar{x}<x_2\leq b$, the proof of the theorem is complete.

REMARK 1. One easily sees that the theorem also holds in case $u(a)=+\infty=u(b)$, or $u(a)=-\infty=u(b)$. In this case, however, the conclusion can be strengthened to $a<x_1<\bar{x}<x_2<b$.

REMARK 2. An identical proof yields a somewhat more general theorem applicable when $u(b)\neq u(a)$. The hypothesis $u(b)=u(a)$ may be replaced by $u(b)=f[u(a)]$ where f is *any* function which is continuous on the range of the function u. The conclusion remains as before, except that the equation $u(x_2)=u(x_1)$ is replaced by $u(x_2)=f[u(x_1)]$.

REMARK 3. If either a or b is regular for (1), the first alternative of the theorem cannot occur. The example (6), with $(a, \bar{x}, b)=(-\infty, 0, \infty)$, $y_1=(1+x^2)^\alpha \times \int_0^x (1+t^2)^{\alpha-1}\, dt$ and $u(x)=(1+x^2)^{-1}$ with $u(\pm\infty)=0$, shows that, in general, the first alternative of the theorem cannot be deleted. I do not know whether the case $\bar{x}_1>\bar{x}$ (or the case $a<t_1<\bar{x}$) can actually occur or not, but this does not affect the statement of the theorem.

References

1. P. R. Beesack, *Integral inequalities of the Wirtinger type*, Duke Math. J. **25** (1958), 477–498.

2. ———, *Integral inequalities involving a function and its derivative*, Amer. Math. Monthly **78** (1971), 705–741.

3. E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, New York, 1955.

4. E. L. Ince, *Ordinary differential equations*, (reprint), Dover, New York, 1944.

Carleton University,
Ottawa, Ontario