Sturm's Separation and Comparison Theorems

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

The prototype equation of this chapter is $y'' + k^2 y = 0$ where k is a positive constant. The solution $C \sin k(x - x_0)$ is oscillatory in $\mathbb{R}$. We observe two features of these solutions. The first is that the zeros of two linearly independent solutions are interlaced. If y_1 and y_2 are two linearly independent solutions and if x_1 and x_2 are two consecutive zeros of y_1, then y_2 has exactly one zero in (x_1, x_2). The second noteworthy feature is that the distance between two consecutive zeros is π/k and becomes smaller as k becomes larger. We generalize these two observations to the case of solutions of differential equations of the form $(ry')'(x) + p(x)y(x) = 0$.

Consider the second order DE

$$a(x)y'' + b(x)y' + c(x)y = 0, \tag{1}$$

where a, b, c are continuous with $a(x) > 0$ on an interval J. If we multiply both sides of the equation (1) by

$$(1/a)e^{\int_{x_0}^{x} \frac{b(x)}{a(x)}\, dx}$$

for $x_0, x \in J$, then the equation (1) can be rewritten in the form

$$[r(x)y']' + p(x)y = 0, \tag{2}$$

where $r > 0$ and r, p are continuous on J. Here,

$$r(x) := e^{\int_{x_0}^{x} \frac{b(x)}{a(x)}\, dx} \quad \text{and} \quad p(x) = \frac{c(x)}{a(x)} e^{\int_{x_0}^{x} \frac{b(x)}{a(x)}\, dx}.$$

Equation (2) is called the *self-adjoint form* of the equation (1).

Example 1. The equations $y'' - y = 0$ and $y'' + y = 0$ are in self-adjoint form. The self-adjoint form of Euler's equation $x^2 y'' - xy' + 2y = 0$ on $(0, \infty)$ is

$$(\frac{1}{x}y')' + \frac{2}{y^3}y = 0.$$

Ex. 2. Put the following DE's in self-adjoint form:
(a) $x^2 y'' + xy' + (x^2 - n^2)y = 0$ (n constant).
(b) $xy'' + (1 - x)y' + ny = 0$ (n constant).
(c) $y'' - 2xy' + 2ny = 0$ (n constant).

Theorem 3 (Sturm Separation Theorem)**.** *Let y_1 and y_2 be two linearly independent solutions of Eq. (2) on the interval $J = [a, b]$. Assume that $r(x) > 0$ on J. Then between any two consecutive zeros of y_1 there will be precisely one zero of y_2.*

Proof. Since y_j are linearly independent, neither can be the zero solution. Let x_1 and x_2 be two consecutive zeros of y_1 with $x_1 < x_2$. Since $-y_1$ is also a solution, we may assume that $y_1(x) > 0$ on (x_1, x_2). Similarly we may assume that $y_2(x_1) \geq 0$. Since y_j's are solutions of the DE (2), we get

$$(ry_1')' + py_1 = 0 \quad \text{and} \quad (ry_2')' + py_2 = 0.$$

Multiplying the first of these equations by y_2 and the second by y_1, and subtracting the resulting equations we obtain

$$y_2(ry_1')' - y_1(ry_2')' = 0 \quad \text{or} \quad y_2(ry_1')' = y_1(ry_2')'$$

Integrate both sides from x_1 to x_2 to get

$$\int_{x_1}^{x_2} y_2(ry_1')'\, dx = \int_{x_1}^{x_2} y_1(ry_2')'\, dx. \tag{3}$$

We integrate both sides of (3) by parts to obtain

$$y_2(ry_1')\,|_{x_1}^{x_2} - \int_{x_1}^{x_2} ry_1'y_2'\, dx = y_1(ry_2')\,|_{x_1}^{x_2} - \int_{x_1}^{x_2} ry_1'y_2'\, dx.$$

Cancelling the equal terms (involving the integrals) and rearranging we have

$$r(x_1)W(y_1, y_2)(x_1) = r(x_2)W(y_1, y_2)(x_2). \tag{4}$$

Thanks to the linear independence of y_1 and y_2, the Wronskian is never zero on J. However, by hypothesis, $y_1(x_1) = 0 = y_1(x_2)$, so that

$$W(x_1) = -y_1'(x_1)y_2(x_1) \neq 0 \quad \text{and} \quad W(x_2) = -y_1'(x_2)y_2(x_2) \neq 0. \tag{5}$$

By our assumption $y_2(x_1) \geq 0$. From (5), we deduce that $y_2(x_1) > 0$ and $y_1'(x_1) \neq 0$. (The latter can be arrived at by the uniqueness and the linear independence also!) Since $y_1(x) > 0$ for $x_1 < x < x_2$, we must have $y_1'(x_1) > 0$. (Why? See Exercise 4 below.) Hence we conclude that $W(x_1) < 0$. Similarly, $y_1'(x_2) < 0$. Since $r(x) > 0$ on J, Eq. (4) implies that $W(x_2)$ is negative. Thus from (5) and the fact that $y_1'(x_2) < 0$, we conclude that $y_2(x_2) < 0$. It follows from the intermediate value theorem that y_2 must have a zero in (x_1, x_2).

Finally, y_2 can have only one zero between x_1 and x_2. If it had more than one, say $x_1 < \alpha < \beta < x_2$ we can reverse the roles of y_1 and y_2 in the above argument to conclude that y_1 must have a zero in $(\alpha, \beta) \subset (x_1, x_2)$. This contradicts our hypothesis that x_1 and x_2 are consecutive zeros of y_1. □

Ex. 4. If u is a solution of (2) such that $u(x_1) = 0 = u(x_2)$ and $u(x) > 0$ on (x_1, x_2), then $u'(x_1) > 0$.

Ex. 5. Prove that between any two consecutive zeros of $\sin x$ there is a zero of $\sin x + \cos x$.

Ex. 6. Show that the functions $f(x) := a\sin x + b\cos x$ and $g(x) := c\sin x + d\cos x$ have alternating zeros whenever $ad - bc \neq 0$.

Ex. 7. Show that the zeros of $\cos\log x$ and $\sin\log x$ alternate.

The Sturm separation theorem indicates that the number of zeros of any two solutions of (2) are approximately the same. However it does not guarantee the existence of any zeros. We now prove a theorem which compares the oscillation of the solutions of two different equations.

Theorem 8 (Sturm Comparison Theorem)**.** *Consider the DE's*

$$(r(x)y')' + p(x)y = 0 \tag{6}$$
$$(r(x)z')' + q(x)z = 0. \tag{7}$$

Assume that $r(x)$ is positive and y and z are solutions of (6) and (7) respectively. Let x_1 and x_2 be two consecutive zeros of y in J. Assume that $q(x) \geq p(x)$ for $x \in J := [a,b]$ with strict inequality holding at some point $x_0 \in (x_1, x_2)$.

If z vanishes at x_1, it will vanish again in (x_1, x_2). Moreover, in that case, every solution of (7) will vanish at some point in the interval (x_1, x_2).

Remark 9. Roughly speaking, the first conclusion says that larger p in (2) the more rapidly its solutions will oscillate.

Proof. As in the last theorem, we may assume without loss of generality, that $y > 0$ on (x_1, x_2) and that $y'(x_1) > 0$ and $y'(x_2) < 0$. Multiplying (6) by $z(x)$ and (7) by $-y(x)$ and adding, we obtain

$$z(ry')' - y(rz')' + (q-p)yz = 0.$$

Integrate this equation from x_1 to x_2 and apply integration by parts to the first two terms. We get

$$\left(zry' \,|_{x_1}^{x_2} - \int_{x_1}^{x_2}\right) - \left(yrw' \,|_{x_1}^{x_2} - \int_{x_1}^{x_2}\right) = \int_{x_1}^{x_2} (q-p)yz\,dx.$$

The two integrals on the left-hand side of this equation cancel out. Gathering like terms yields

$$r(x_2)W(z,y)(x_2) - r(x_1)W(z,y)(x_1) = \int_{x_1}^{x_2} (q-p)yz\,dx. \tag{8}$$

Using the facts $z(x_1) = 0 = y(x_1)$ and $y(x_2) = 0$ the equation (8) reduces to

$$r(x_2)z(x_2)y'(x_2) = \int_{x_1}^{x_2} (q-p)yz\,dx. \tag{9}$$

If $z(x) > 0$ on (x_1, x_2), then the integrand on the RHS of (9) is nonnegative. It is, in fact, positive, since $q(x_0) > p(x_0)$ for some $x_0 \in (x_1, x_2)$. (Use the continuity of $q - p$ and the monotonicity of the integral.) We therefore conclude that the RHS of (9) is positive. Since $r(x_2) > 0$ and $y'(x_2) < 0$, it follows that $z(x_2) < 0$. Since $z > 0$ on (x_1, x_2), it follows by continuity that $z(x_2) \geq 0$. It follows from this contradiction that z must vanish at some point in (x_1, x_2).

If z_1 is another solution of (7), not a scalar multiple of z, then by Sturm's separation theorem, z_1 must have a zero between the zeros of z and therefore in the interval (x_1, x_2). □

Corollary 10. *Consider the equation $y'' + py = 0$. Assume that $0 \le m \le p(x) \le M$ for $x \in (a, b)$. If a solution of this equation has two consecutive zeros $x_1 < x_2$ in (a, b), then*

$$M^{-1/2}\pi \le x_2 - x_1 \le m^{-\pi/2}m.$$

Proof. Compare this equation with $z'' + Mz = 0$. Then $z(x) = \sin\sqrt{M}(x - x_1)$ is a solution with a zero at x_1. The zeros of this solution nearest to x_1 are at a distance $M^{-1/2}\pi$, the first inequality follows. The second is proved in an analogous manner by considering $z'' + my = 0$. □

Example 11 (A Typical Application: Bessel's Equation)**.** Consider the Bessel's equation of order p

$$x^2y'' + xy' + (x^2 - p^2)y = 0, (x \in (0, \infty)),$$

where p is a constant. This can be transformed by the substitution $y = u/\sqrt{x}$ to

$$u'' + \left(1 + \frac{1 - 4p^2}{4x^2}\right) u = 0. \tag{10}$$

(Note that this substitution does not change the zeros!) We compare this with $z'' + z = 0$ and arrive at the following conclusions.

(a) For $0 \le p < 1/2$, every solution of Bessel's equation of order p has at least one zero in every interval $\subset (0, \infty)$ of length π.

(b) If $p = 1/2$, every nontrivial solution of Bessel's equation of order p,

$$y = u/\sqrt{x} = \frac{1}{\sqrt{x}}(c_1 \cos x + c_2 \sin x),$$

has zeros separated by an interval of length π for $x > 0$.

(c) If $p > 1/2$, every solution of Bessel's equation of order $p$can have *at most* one zero in every interval of length π.

Ex. 12. Let $p < 0$ on $[a, b]$. Let y be a nontrivial solution of the equation $y'' + py = 0$. Then y can have at most one zero in $[a, b]$. *Hint:* Suppose not. Let x_1 and x_2 with $a \le x_1 < x_2 \le b$ be zeros of y. Compare this solution with the solution $z = 1$ of $z'' = 0$.

Ex. 13. With the notation of Cor. 10 show that the number n of zeros in $(x_1, x) \subset (a, b)$ satisfies the following inequalities

$$\frac{x - x_1}{\pi}\sqrt{m} < n < \frac{x - x_1}{\pi}\sqrt{M}.$$

Ex. 14. Show that every solution of $y'' + t^2y = 0$ on $[1, \infty)$ must have a zero in $[1, \infty)$. *Hint:* Compare with $z'' + z = 0$.

Ex. 15. Which of the equations (a) $y'' + (1 + x^2)y = 0$, (b) $z'' + 2xz = 0$ has the more rapidly oscillating solution in the interval $[1, 10]$?

Ex. 16. How many zeros does every solution of $y'' + xy = 0$ have in the interval $(0, \infty)$?

Ex. 17. Prove that every solution of Bessel's equation of order $p \geq 0$ has infinitely many zeros in $(0, \infty)$.

We end this topic with a simple proof of the following well-known oscillation theorem.

Theorem 18. *Consider* $(ry')' + py = 0$ *where* $r > 0$ *and* r, p *are continuous. Assume that* $\int^\infty (1/r) = \infty = \int^\infty p$*, where the lower limit is arbitrarily large. Then any solution of the equation oscillates in* $(0, \infty)$*.*

Proof. Assume the contrary. Then there exists a solution y and $A > 0$ such that y has no zeros in (A, ∞). Let $z := (ry')/y$. Then z is a solution of the (Riccati) equation $z' + (z^2/r) + p = 0$ on (A, ∞). Integrating this equation between A and t sufficiently large yields

$$z(t) + \int_A^t \frac{z^2}{r} = z(a) - \int_A^\infty p < 0. \tag{11}$$

If we let $R(t) := \int_A^t \frac{z^2}{r}$, then (11) tells us that $R^2 \leq R'r$. Separating the variables and integrating we get

$$\int_A^t \frac{1}{r} = \int_A^t \frac{dR}{R^2} = \frac{1}{R(A)} - \frac{1}{R(t)} \leq \frac{1}{R(A)}$$

which contradicts our hypothesis that $\int^\infty (1/r) = \infty$. □

Ex. 19. Give another solution of Ex. 17.

Ex. 20. Let y be a nonzero solution of $y'' + py = 0$ on $J := [a, b]$. Show that y cannot have infinitely many zeros in J.