

## CANONICAL AND HAMILTONIAN FORMALISM APPLIED TO THE STURM-LIOUVILLE EQUATION\*

BY

M. A. BIOT (*Shell Development Company, N. Y.*)

AND

I. TOLSTOY (*Columbia University, Hudson Laboratories, Dobbs Ferry, New York*)

**Abstract.** The Sturm-Liouville equation is expressed in Hamiltonian form. A simple generating function is derived which defines a large class of canonical transformations and reduces the Sturm-Liouville equation to the solution of a first order equation with a single unknown. The method is developed with particular reference to the wave equation. The procedure unifies many apparently diverse treatments and leads to new insights and procedures. Some new transformations are obtained, useful in the turning point region and for the improvement of accuracy in the region of validity of W.K.B. solutions. In addition a new power series expansion near the turning point is obtained.

**1. Standard canonical form of the Sturm-Liouville equation.** Let  $q, p$  be conjugate variables and  $f, g$  arbitrary functions of the independent coordinate  $z$ . The Hamiltonian:

$$H = \frac{1}{2} \left[ \frac{1}{f} p^2 + gq^2 \right] \quad (1.1)$$

defines the canonical equations

$$\frac{dq}{dz} = \frac{1}{f} p \quad (1.2)$$

$$\frac{dp}{dz} = -gq$$

which are equivalent to the second order equation

$$\frac{d}{dz} \left( f \frac{dq}{dz} \right) + gq = 0, \quad (1.3)$$

i.e., the general Sturm-Liouville equation, having a vast number of physical applications. A very large fraction of the integrable equations of mathematical physics are of this type. Thus, in separable coordinate systems, the solution of the partial differential equations of Laplace, of heat flow, of mechanical electro-magnetic and quantum-mechanical waves reduces to some form of Eq. (1.3). However, the exact behavior of the solutions of this equation is known only for a relatively few simple cases. In problems of practical interest it occurs frequently that the functional forms  $f$  and  $g$  are such that no solutions of Eq. (1.3) in terms of known functions is possible. One is thus often reduced to devising approximations and/or numerical schemes of integration.

Hamilton's equations (1.2) are a system of first-order equations entirely equivalent to the Sturm-Liouville equation (1.3). From a purely numerical standpoint one might expect that more advantageous formulations are possible and, since we are dealing with

---

\*Received August 17, 1959.

a set of canonical equations, it seems only natural to apply the theory of canonical transformations which provides a natural and flexible means of exploring equivalent systems of first order equations. In the following sections we will describe one group of transformations, which should not only have certain numerical advantages, but also provide new insights and results of a general nature.

**2. Canonical transformation of the standard form.** Clearly, considerable advantages can accrue from finding an equivalent set of first order equations, one of which is uncoupled, the other being soluble by quadrature. These requirements may be satisfied if one defines equivalent canonical equations in the new variables corresponding to the following class of generating functions:

$$F(q, P, z) = q^2 \Phi(P, z), \quad (2.1)$$

where  $\Phi$  is an arbitrary function. The corresponding transformation is:

$$p = \frac{\partial F}{\partial q} = 2q\Phi \quad (2.2)$$

$$Q = \frac{\partial F}{\partial P} = q^2 \frac{\partial \Phi}{\partial P} \quad (2.3)$$

and

$$q^2 = \frac{Q}{\frac{\partial \Phi}{\partial P}} \quad (2.4)$$

$$p^2 = 4 \frac{Q\Phi^2}{\frac{\partial \Phi}{\partial P}}. \quad (2.5)$$

The new Hamiltonian:

$$H = \frac{1}{2} \left[ \frac{1}{f} p^2 + gq^2 \right] + \frac{\partial F}{\partial z} = \frac{1}{2} \left[ \frac{1}{f} p^2 + gq^2 \right] + q^2 \frac{\partial \Phi}{\partial z} \quad (2.6)$$

becomes, after substituting (2.4), (2.5):

$$H = QK(P, z) \quad (2.7)$$

with

$$K(P, z) = \frac{1}{2} \left[ \frac{4\Phi^2}{f \frac{\partial \Phi}{\partial P}} + \frac{g}{\frac{\partial \Phi}{\partial P}} \right] + \frac{\frac{\partial \Phi}{\partial z}}{\frac{\partial \Phi}{\partial P}}. \quad (2.8)$$

The canonical equations for  $Q, P$  are

$$\frac{dP}{dz} = -K, \quad (2.9)$$

$$\frac{dQ}{dz} = Q \frac{\partial K}{\partial P}. \quad (2.10)$$

The last is integrable by quadrature:

$$Q = Q_0 \exp \left[ \int_{z_0}^z \frac{\partial K}{\partial P} dz \right] \quad (2.11)$$

**3. Selection of transformations.** Assume that

$$\phi(P, z) = \varphi(z)\psi(P) \quad (3.1)$$

then

$$K = \frac{1}{2} \left[ \frac{4\varphi\psi^2}{f \frac{d\psi}{dP}} + \frac{g}{\varphi \frac{d\psi}{dP}} \right] + \frac{\frac{d}{dz} \text{Log } \varphi}{\frac{d}{dP} \text{Log } \psi}. \quad (3.2)$$

One may write, without loss of generality:

$$\psi = \frac{S(P)}{R(P)}. \quad (3.3)$$

This facilitates the selection of the  $Q, P$  that are most desirable for a given problem or class of problems, since now

$$q = Q^{1/2}R(P) \bar{\varphi}^{-1/2} \Delta^{-1/2} \quad (3.4)$$

$$p = 2Q^{1/2}S(P) \varphi^{1/2} \Delta^{-1/2}, \quad (3.5)$$

with

$$\Delta = R \frac{dS}{dP} - S \frac{dR}{dP}. \quad (3.6)$$

In this case:

$$K = \frac{1}{\Delta} \left\{ \frac{1}{2} \left[ \frac{4}{f} \varphi S^2 + \frac{g}{\varphi} R^2 \right] + RS \frac{d}{dz} \text{Log } \varphi \right\}. \quad (3.7)$$

Equations (3.4), (3.5) enable one to select explicitly the forms  $R(P), S(P)$  in an appropriate fashion, since  $R, S$  are arbitrary. This approach gives a free rein to physical intuition. Generally speaking, in the selection of transformations one may adopt one of two attitudes.

On one hand, one can look for special properties that would be advantageous for a specific problem. Thus, if one is in a position to anticipate good formal approximations to  $q, p$  the functions  $R, S$ , can be chosen so as to produce the corresponding forms in Eqs. (3.4), (3.5). This will result in almost linear or almost constant solutions to Eqs. (2.9), (2.10) and thus greatly facilitate numerical integration. But this approach is fruitful in a relatively small number of cases and will not be discussed here. Besides the various perturbation schemes, such as that of Miller and Good [1] are well suited to this kind of situation.

On the other hand, one may wish for properties suitable for integration in many cases, so as to provide a generally useful method; this is essentially the point of view adopted in the following sections, and which we will illustrate in the context of the wave equation.

Note also that the choice of transformations need not be based upon Eqs. (3.4), (3.5), but can stem from formal considerations concerning the desired behavior or form of individual terms in  $K(P, z)$ , Eq. (3.2). This approach is less intuitive than the one we have followed. We shall have occasion to remark further upon some of these transformations.

**4. The wave equation: a polar type of transformation.** If in Eq. (1.3) we take

$$\begin{aligned} f &= 1, \\ g &= \beta^2 = h(z) - k^2, \end{aligned} \quad (4.1)$$

where  $k$  is a parameter, we have the time-independent wave equation

$$\frac{d^2 q}{dz^2} + \beta^2 q = 0. \quad (4.2)$$

This equation is typical of equations with oscillatory solutions and thus provides particular motivation for the use of transformations of the class discussed in the preceding section. For example, one may look for functions  $R(P)$ ,  $S(P)$ , satisfying the following criteria:

- (1) The unknowns  $P$ ,  $Q$  are to be reasonably monotonic while representing at the same time functions  $q$ ,  $p$  that may be highly oscillatory.
- (2) Neither of these functions, nor any of the coefficients of their defining equations shall be singular in the region of interest.

Criterion (1) is readily fulfilled by taking for  $R$ ,  $S$  forms such that their dependence upon  $P$  is of the same general type as the expected dependence of  $q$  on  $z$ . This can be achieved approximately, in a broad and general fashion, by taking for  $R$ ,  $S$  trigonometric, exponential or Bessel functions of  $P$ .

Criteria (2) require only that  $R$ ,  $S$ ,  $\varphi$  be functions without singularities and that  $\Delta$ ,  $\varphi$  have no zeros.

A simple and obvious series of choices meet these requirements. First of all, let

$$\varphi = \text{const} = 2^{-1/2}. \quad (4.3)$$

Note that this corresponds to dropping the last term in  $K(P, z)$ , Eq. (3.2), thus greatly simplifying the formal problem.

Next, take

$$\begin{aligned} R &= \sin P, \\ S &= -\cos P, \end{aligned} \quad (4.4)$$

this gives

$$\begin{aligned} q &= (2Q)^{1/2} \sin P, \\ p &= -(2Q)^{1/2} \cos P. \end{aligned} \quad (4.5)$$

And by virtue of Eqs. (2.9), (3.7),  $P$  obeys the equation

$$\frac{dP}{dz} + \cos^2 P + \beta^2 \sin^2 P = 0 \quad (4.6)$$

and

$$Q = Q_0 \exp \left[ \int_{z_0}^{z^*} (\beta^2 - 1) \sin 2P \, dz \right]. \quad (4.7)$$

$QP$ , and the coefficients of the differential equation (4.6) are well behaved everywhere, including the region of the turning point  $\beta^2 = 0$ . Boundary conditions at a discontinuity imply simply continuity of  $P$ . This formulation has therefore definite merits of simplicity in appearance and in analytical behavior, as well as generality. In this sense, it has obvious advantages over other, better known, equivalent systems of first order equations sometimes used in practice, such as the Riccati equations for the reflection coefficient or for the impedance: the former has oscillatory solutions and a coefficient that is singular at turning points, and the latter has solutions that are both oscillatory and singular.

The transformation (4.5) is, of course, a particular case of the more general group for which  $\varphi$  is not constant. Because of their aspect, it seems natural to call them polar-type transformations. The corresponding generating function

$$F = -q^2 \varphi \cotan P \quad (4.8)$$

is the generalization to time-dependent systems of the usual function for oscillating systems found in many textbooks. Obviously, different choices of  $\varphi$  may result in transformations that would be better adapted to specific cases. But, of the transformations satisfying the second criterion of analyticity, Eqs. (4.5) give one of the more general types since, because of its relatively non-committal form, it has a wide range of application.

**5. A polar transformation for slowly variable parameters: Relation to the W. K. B. approximation.** It may sometimes be expedient to ignore criteria (2) of the preceding section and to permit the appearance of singular coefficients, providing that the singularities appear outside of the range of intended applications of the resulting equations. The following purely physical reasoning leads to a useful transformation of that type.

One may interpret Eq. (4.2) as the equation for a pendulum of variable length, of period  $2\pi/\beta$ . The average energy per cycle is

$$\bar{E} = \overline{\beta^2 q^2}. \quad (5.1)$$

If  $\beta$  varies very slowly one has, by *Ehrenfest's theorem* for slowly varying constraints, the adiabatic invariant

$$\frac{\tau\pi}{\beta} \bar{E} = \text{const.} \quad (5.2)$$

And, since  $\beta$  varies slowly, for each cycle:

$$\bar{q}^2 = \frac{1}{\beta} \times \text{const.} \quad (5.3)$$

The constant term arises from the time average over a cycle of an oscillating, almost periodic function, i.e.,  $q$  is of the form

$$q = \beta^{-1/2} f(z), \quad (5.4)$$

where  $f(z)$  is an almost periodic function of period  $2\pi/\beta$ . It is therefore only natural to choose  $\varphi$  proportional to  $\beta$  in Eqs. (3.4), (3.5) e.g.,

$$\varphi = \frac{1}{2}\beta. \quad (5.5)$$

Using script capitals for our new canonical variables  $\mathcal{Q}$   $\mathcal{P}$  to distinguish them from the  $Q$ ,  $P$  of Eqs. (4.5), we have

$$\begin{aligned} q &= \left(\frac{2\mathcal{Q}}{\beta}\right)^{1/2} \sin \mathcal{P}, \\ p &= -(2\mathcal{Q}\beta)^{1/2} \cos \mathcal{P}, \end{aligned} \quad (5.6)$$

corresponding to the generating function

$$F = -\frac{1}{2}q^2\beta \cotan \mathcal{P}. \quad (5.7)$$

This can also be obtained by imposing formal requirements on  $K(P, z)$  in Eq. (3.2). Thus, with  $f = 1$  and Eq. (5.5) one has

$$K = \pm\beta \left[ \frac{\psi^2}{\frac{d\psi}{d\mathcal{P}}} + \frac{1}{\frac{d\psi}{d\mathcal{P}}} \right] + \frac{\psi}{\frac{d\psi}{d\mathcal{P}}} \frac{d}{dz} \text{Log } \beta. \quad (5.8)$$

Choosing the  $-$  sign in front of the bracket leads to Eq. (5.7) if one imposes the condition that the term in brackets be equal to  $-1$ . The equations for  $\mathcal{P}$ ,  $\mathcal{Q}$  are

$$\frac{d\mathcal{P}}{dz} - \frac{1}{2} \frac{d}{dz} (\text{Log } \beta) \sin 2\mathcal{P} + \beta = 0, \quad (5.9)$$

$$\mathcal{Q} = \mathcal{Q}_0 \exp \left[ - \int_{z_0}^z \cos 2\mathcal{P} d \text{Log } \beta \right]. \quad (5.10)$$

An instructive change of independent variable is

$$s = - \int_{z_0}^z \beta dz, \quad (5.11)$$

giving

$$\frac{d\mathcal{P}}{ds} - \frac{1}{2} \frac{d}{ds} (\text{Log } \beta) \sin 2\mathcal{P} - 1 = 0. \quad (5.12)$$

In regions of very slowly varying  $\beta$

$$\frac{d}{ds} \text{Log } \beta \ll 1 \quad (5.13)$$

an approximate solution of Eq. (5.12) is

$$\mathcal{P}(s) \approx s + \mathcal{P}_0, \quad (5.14)$$

i.e.,  $\mathcal{P}(s)$  becomes a straight line of slope unity. Condition (5.13) can be interpreted physically: it implies that the change in parameters is so small that the backscattering or continuous process of partial reflection of a wave train can be neglected [2]. This backscatter and multiple scatter is measured by

$$\frac{d}{ds} \log \beta = \frac{1}{\beta} \frac{d\beta}{ds}.$$

It will be small if either  $d\beta/ds$  is small or if  $\beta$  is large, (small wave length). Elsewhere  $\mathcal{P}$

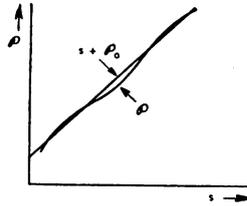


FIG. 1. Comparison of exact and approximate solutions of Eq. (5.12).

will generally be a monotonic wave, gently oscillating about the straight line (5.14), as shown in Fig. 1.

$$Q \cong Q_0 = \text{const} \tag{5.15}$$

and

$$q = A\beta^{-1/2} \sin(s + \phi_0), \tag{5.16}$$

where  $A$  is a constant. This is the first-order W.K.B. approximation.

We thus have a new interpretation of the W.K.B. formulae: the phase term is the limiting form for slowly variable parameters of our canonical  $Q$  variable, whereas the constant of integration  $A$  is the limit of  $(2Q)^{1/2}$  under the same conditions.

In the vicinity of a turning point  $\beta = 0$  the criterion (5.13) is no longer fulfilled: the variation of parameters per unit  $z$  wave length becomes very fast and the reflected components become large (since one is in the region of glancing incidence and total reflection). Indeed, the coefficient:

$$\frac{d}{ds} \text{Log } \beta = \frac{1}{2} \frac{dh}{dz} \frac{1}{\beta^2} \tag{5.17}$$

is divergent. As could have been foreseen: the representation (5.6) is not felicitous for rapidly varying parameters since we had deliberately designed it for slowly varying  $\beta$ .

Here, the transformation of Sec. 4 is much more appropriate. It is connected to our  $Q$  variable thus:

$$P = \tan^{-1} \left[ \frac{1}{\beta} \tan Q \right]. \tag{5.18}$$

The behavior of  $Q$  and  $P$  far from a turning point is similar only for  $\beta \approx 1$ . Otherwise  $P$  consists of a sequence of steps crossing back and forth through the  $Q$  curve (Fig. 2).

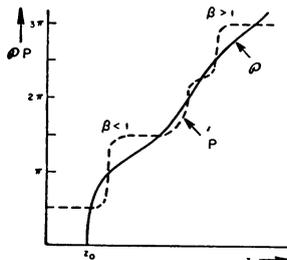


FIG. 2. General behavior of variables  $P$  and  $Q$ .

**6. A new turning-point expansion.** An expansion of  $P$ , Eq. (4.6), in power series of  $\beta^2$  is easily achieved, leading to a general result valid on both sides of the turning point  $\beta^2 = 0$ , for any functional dependence of  $\beta^2$  on  $z$ .

For example, we may write

$$\tan P = a_0 + a_2\beta^2 + a_4\beta^4 + \dots, \quad (6.1)$$

where  $a_0$  is an integration constant. Substitution in Eq. (4.6) yields, writing  $dh/dz = h'$ :

$$\begin{aligned} a_2 &= -\frac{1}{h'}, \\ a_4 &= -\frac{1}{2h'}(a_0^2 + a_2^2), \\ a_6 &= -\frac{1}{3h'}(2a_0a_2 + a_2^2a_4). \end{aligned} \quad (6.2)$$

As pointed out below  $\tan P$  is the impedance. This gives a physical interpretation to the expansion (6.1). It is also easily verified that  $\tan P$  satisfies a Ricatti equation. Applying Eqs. (4.5) and (4.7) and neglecting terms of the order of  $\beta^4$  and higher gives:

$$q = A \left[ a_0 - \frac{1}{h'}\beta^2 + \frac{1}{2}a_0^3\beta^2 \right] \exp \left[ a_0 \frac{\beta^2}{2} + \epsilon(\beta^4) \right]. \quad (6.3)$$

To extend this to terms of the order of  $a_0^4\beta^4$  and higher, take

$$\epsilon(\beta^4) = -\frac{1}{4}a_0^4\beta^4 - a_0 \int^z \beta^2 dz. \quad (6.4)$$

Additional terms are easily secured. These expansions are convergent in the vicinity of  $\beta = 0$ , but become divergent in the region  $P > \tan^{-1} 1/\beta$ . We have not made a rigorous study of the radius of convergence of the series (6.1).

For a small region near a simple turning point ( $h' \neq 0$ ) we can approximate  $\beta^2$  by a linear function of  $z - z_0$ ,  $z_0$  being the turning-point coordinate. Equation (6.3) then gives the wave function as a constant plus a term linear in  $(z - z_0)$ . This same result is obtained by taking Langer's approximation [3], which is of the type:

$$y = (z - z_0)^{1/2} [BJ_{1/3}\{(z - z_0)^{3/2}\} + CJ_{-1/3}\{(z - z_0)^{3/2}\}], \quad (6.5)$$

expanding it in powers of  $(z - z_0)$  and keeping the resulting constant and linear term. Our results (6.1)–(6.3) are more general since they are the correct expansions for any functional dependence of  $\beta$  upon  $z$ , whereas Langer's formulae are based upon the approximation that  $\beta^2$  is linear in a small region surrounding the turning point.

**7. General remarks.** The discussion of Secs. 4, 5, 6 would not be complete without some brief remarks concerning the physical meaning and previous derivations and uses of the  $P$  and  $\mathcal{O}$  variables in the context of the wave equation.

The meaning of  $P$  is obvious from the defining equations (4.5):

$$\frac{q}{q'} = \zeta = \tan P \quad (7.1)$$

is an impedance. The use of  $P$  is therefore equivalent to replacing impedances  $\zeta$  by  $\tan^{-1} \zeta$  thus eliminating the singularities present in impedance solutions.

As for  $\mathcal{O}$ , a simple physical interpretation in the case of total reflection can be obtained as follows. Consider the well-known Riccati equation for the reflection coefficient  $\mathcal{R}$  [14]

$$\frac{d\mathcal{R}}{dz} - \frac{1}{2} \frac{d}{dz} \text{Log } \beta(1 - \mathcal{R}^2) + 2i\beta\mathcal{R} = 0 \quad (7.2)$$

the substitution

$$\mathcal{R} = e^{2i\varphi} \quad (7.3)$$

shows that  $\varphi$  obeys Eq. (5.9). In other words, when  $\varphi$  is real (total reflection), it is identical to  $\mathcal{O}$ , which is thus the half-change in phase of a plane, totally reflected wave.

The transformation (7.3) has been used by a number of authors, notably Eckersley [4], Hartree [5], Bremmer [6], Brekhovskih [7] and others for approximate descriptions of wave propagation in stratified media. Walker and Wax [8] suggested the use of Eqs. (7.3) and (5.9) for the numerical integration of transmission line problems. Tolstoy [9] has used exact recursive relations for  $\mathcal{O}$  in layered media when  $\text{Log } \beta$  is a sum of step functions, relations that are very similar to those used by Biot [10] in studying the torsional oscillations of loaded rods.

A transformation which is not canonical but very similar to (5.6)

$$\begin{aligned} q &= \rho \sin \theta, \\ q' &= \rho\beta \cos \theta, \end{aligned} \quad (7.4)$$

was first introduced by Prüfer [11], and used again later by Barrett [12] in investigating asymptotic properties of Sturm-Liouville operations. Atkinson [13] pointed out its possible applications to numerical problems. Its usefulness in that sense would be about equivalent to (5.6).

**8. Conclusions.** We have presented a class of canonical transformations which reduces the Sturm-Liouville equation to two first order equations having a number of interesting properties: an equation in  $P$  which is independent of  $Q$ , and an equation in  $Q$  which is soluble by quadrature. The general form of these transformations is quite flexible and involves arbitrary functions of  $P$ ,  $z$ . In particular it is easy to generate transformations of the dependent variable into any desired functional forms, e.g., forms having properties considered suitable for a given problem or class of problems. This leads to new insights and forms useful for general numerical integration schemes.

We have contented ourselves here with illustrating the potential of this approach by simple examples in the context of the wave-equation. In the process we have been led to several interesting results: first, a new transformation (Sec. 4) suitable for numerical integration for all values of the parameters, valid on both sides of a turning point. A second result is a transformation mostly useful for slowly varying parameters; when this variation is slow enough it leads to the W.K.B. results, thus giving a new interpretation of these formulae in terms of canonical variables. Thirdly, a power series expansion in terms of the wave number  $\beta$  leads to a novel and general formula for the wave function near, and to both sides of, a turning point, valid for any functional dependence of the coordinate.

## REFERENCES

1. S. C. Miller and R. H. Good, *Phys. Rev.* **91**, 174 (1953)
2. H. Bremmer, *Communs. Pure and Appl. Math.* **IV**, 1, 103 (1951)
3. R. E. Langer, *Phys. Rev.* **51**, 669 (1937)
4. T. L. Eckersley, *Proc. Roy. Soc.* **A132**, 83 (1931)
5. D. R. Hartee, *Proc. Roy. Soc.* **A131**, 428 (1931)
6. H. Bremmer, *Terrestrial radio waves*, Elsevier, N. Y., 1949
7. L. M. Brekhovskih, *Zhur. Tekh. Fiz.* **XVIII** **4**, 455 (1948); *Izvest. Akad. Nauk., Ser. Fiz* **XIII** **5**, 505 (1949)
8. L. R. Walker and N. Wax, *J. Appl. Phys.* **17**, 1043 (1946)
9. I. Tolstoy, *J. Acoust. Soc. Am.* **27**, 274 (1955); **27**, 897 (1955)
10. M. A. Biot, *J. Appl. Phys.* **11**, 530 (1940)
11. H. Prüfer, *Math. Ann.* **94**, 498 (1928)
12. J. H. Barrett, *Proc. Am. Math. Soc.* **6**, 247 (1955)
13. F. V. Atkinson, *Univ. Nac. Tucumah, Revista A*, **8**, 71 (1951)
14. S. A. Schelkunoff, *Communs. Pure and Appl. Math.* **4**, 1, 117 (1951)