

Complete Solution of a General Problem of Three Bodies

VICTOR SZEBEHELY AND C. FREDERICK PETERS
Yale University Observatory, New Haven, Connecticut
 (Received 16 May 1967)

A problem investigated by C. Burrau in 1913 which may be referred to as the Pythagorean problem of three bodies is solved numerically, the solution being valid for any length of time. The solution found is neither quasi-periodic nor periodic but it assumes the form known in the recent Soviet literature as "elliptic-hyperbolic." In the final configuration two of the three participating bodies form a permanent binary while the third body is rejected to infinity. A new method of treating close approaches which allows achieving the solution is also described.

INTRODUCTION

BY complete solution we mean the description of the motion of the participating three bodies for any length of time. This may be accomplished for periodic, almost periodic, or asymptotic solutions.

Burrau's original paper of 1913 computes the development of this dynamical system for a time that might be considered to be too short to arrive at general conclusions. In the past year, at the suggestion of one of us (Szebehely), research has been conducted at several institutes to "complete" Burrau's problem. At Yale University, Mr. E. M. Standish (1966), under the direction of the first author; at the Institute of Space Studies, NASA, New York, Dr. R. Spinelli (1966) with the cooperation of Dr. M. Lecar and Dr. V. Szebehely; and at the Eidgenössische Technische Hochschule in Zürich Mr. L. Stanek (1966), under the supervision of Dr. E. Stiefel, have studied the problem. The final and complete solution has been obtained only recently, after a new method has been introduced by one of us (Peters) to integrate the equations of motion. The development of this new method was necessary in order to obtain the complete solution because of the occurrence of repeated close approaches during the evolution of the system.

The asymptotic solution found by us may be termed elliptic-hyperbolic. Such final configurations have been predicted by Merman (1958) and by Alexeev (1961). Two of the three bodies move on approximately elliptic orbits around each other forming a close binary system and the third body departs on a hyperbolic orbit to infinity. The total energy of the system is negative.

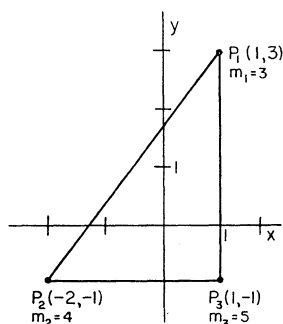


FIG. 1. Initial configuration of the Pythagorean problem of three bodies.

The center of mass of the system is stationary and consequently the close binary and the escaping body move to infinity in opposite directions. The solution presented, therefore, may be considered an example of escape and also of capture, since two of the bodies stay permanently together, while the third is ejected. Other special features of the motion are pointed out in the next section.

A detailed description of the historical and scientific perspective of Burrau's problem has been given recently by the first author (1967).

EVOLUTION OF THE DYNAMICAL SYSTEM

The initial conditions are such that a Pythagorean triangle is described at $t=0$ by the participating bodies. Burrau's example is Pythagorean not only in the geometric sense but also regarding the masses. The sides of the triangle formed by the three bodies at $t=0$ are 3, 4, and 5 and the masses of the participating bodies are also 3, 4, and 5. The particle with mass $m_1=3$, referred to as the first body, is located at that apex of the triangle which is opposite to the side with length 3. The second body with mass $m_2=4$ is positioned opposite to the side whose length is 4 and the third body with $m_3=5$ occupies the third apex. Figure 1 shows the initial configuration and the coordinate system used. The three initial speeds are zero, consequently the motion is planar. The three components of the vector of the angular momentum are zero initially and they remain zero during the motion. The center of mass is initially at the origin of the coordinate system selected and it stays there during the evolution of the system. The total energy E of the system is also preserved and it is identical with the potential energy V at $t=0$. Its value is $E=V=769/60=-12.8166\dots$ using unity for the constant of gravitation.

The development of the system is shown in Figs. 2 to 8. In these figures the motion of the first body is shown by a dotted line, the dashes represent the orbit of m_2 and the solid line illustrates the motion of the body with the largest mass m_3 . The numbers on the curves are the values of the time in unit increments, excepting Fig. 7, where intermediate time steps are also shown. All orbits are referred to the inertial rectangular coordinate system described previously.

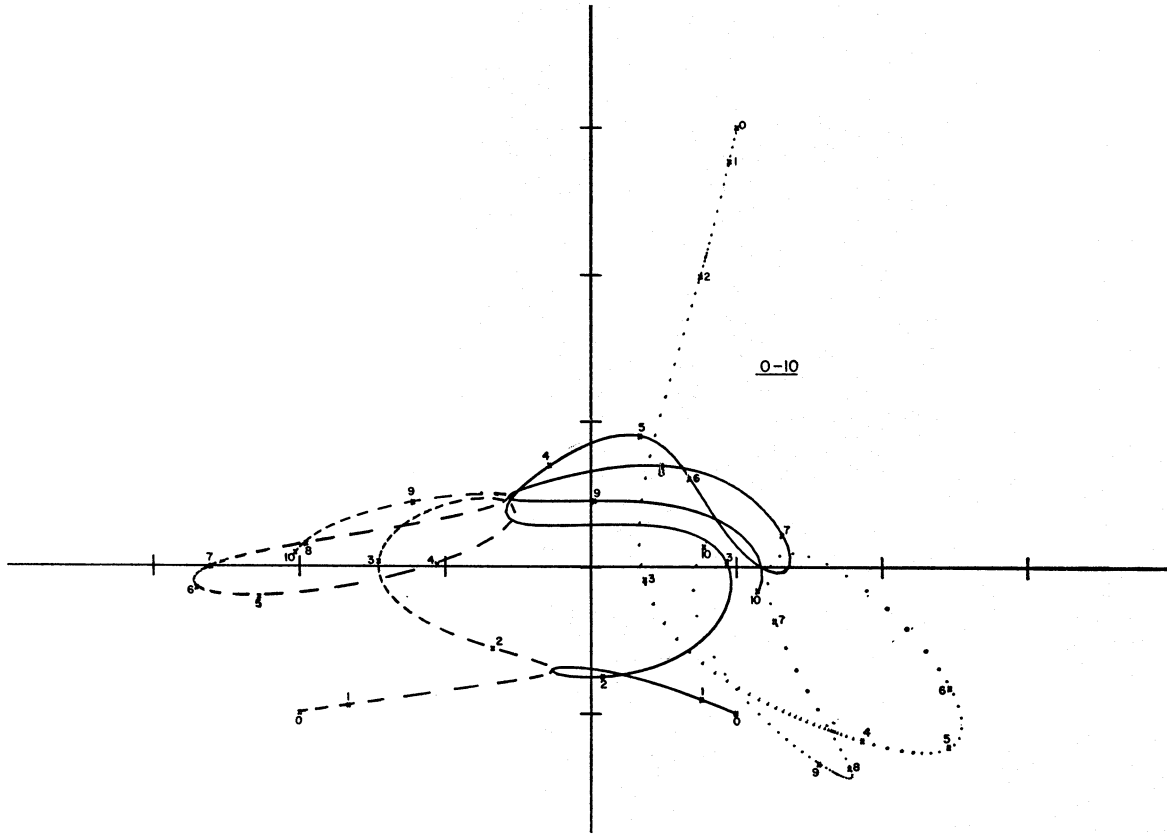


FIG. 2. Orbits between $t=0$ and $t=10$.

The marks on the horizontal x axis and on the vertical y axis are at unit distances. The series of seven figures (2 to 8) were prepared by an automatic plotter at the Yale University Computer Center where the numerical integrations were also performed.

To give physical meaning to the numerical values of the time variable shown on the figures, we note that the equations of motion are integrated using unity for the dimensionless parameter of the problem; i.e.,

$$T^2GM/L^3=1.$$

Here T , L , and M are the time, length, and mass in the system of units used and $G=6.67 \times 10^{-8}$ is the constant of gravitation in cgs units. The preceding equation follows immediately from the equations of motion, which are written as

$$\frac{d^2\mathbf{r}_i}{dt^2} = G \sum_{\substack{j=1 \\ j \neq i}}^3 \mu_j \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3}; \quad i = 1, 2, 3,$$

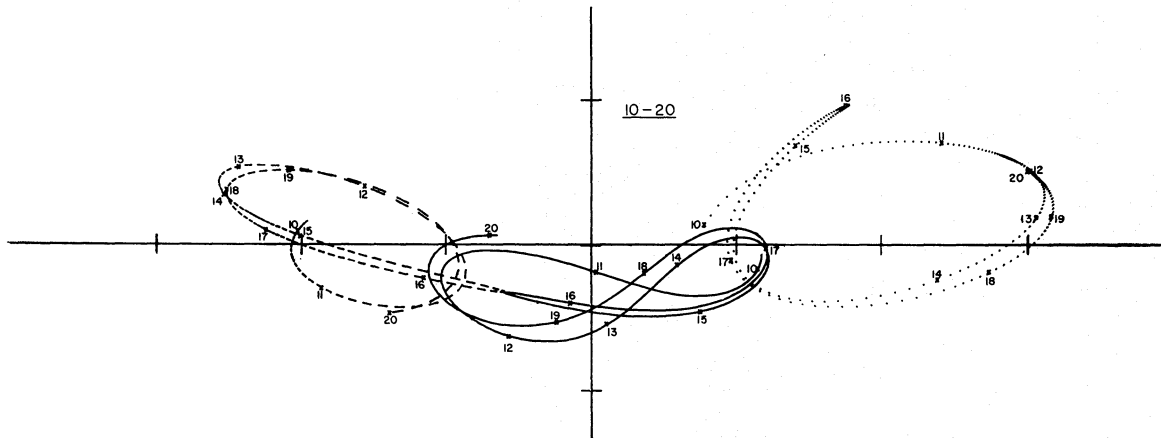
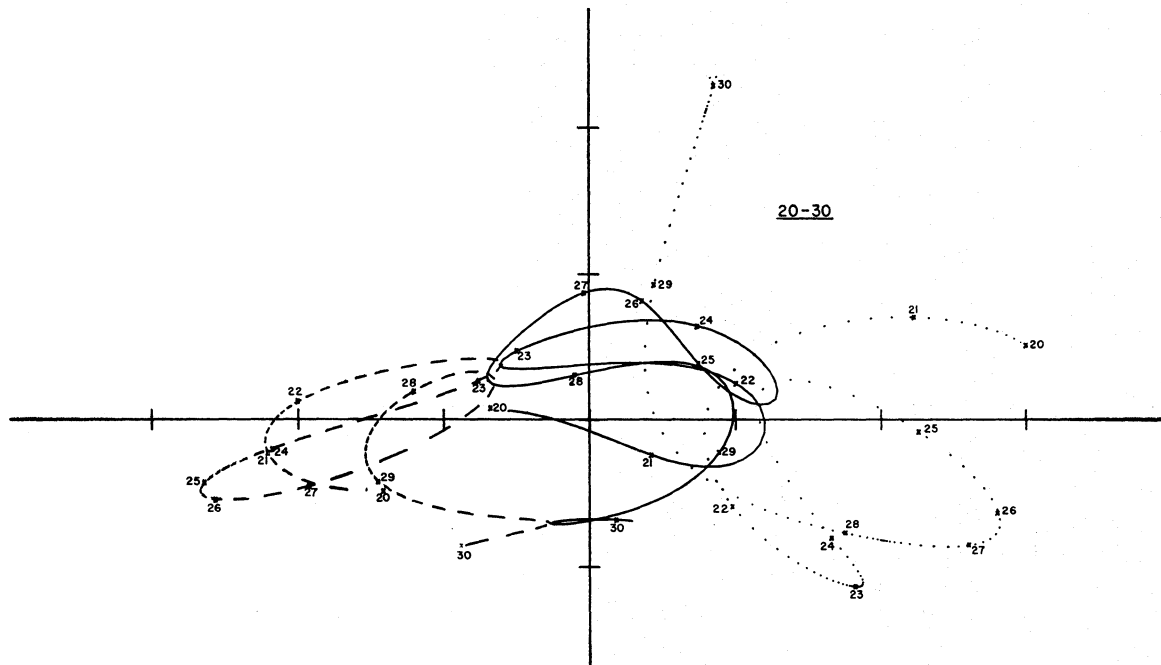
where \mathbf{r}_i is the position vector and μ_i is the mass of the i th body. The dimensionless form of this equation is

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{GMT^2}{L^3} \sum_{\substack{j=1 \\ j \neq i}}^3 m_j \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3},$$

where $\mathbf{r}_i = \mathbf{r}_i/L$, $m_i = \mu_i/M$, and $t = \tau/T$.

Expressing the masses in grams and the distances in centimeters, for instance, the unit of time becomes 3872 sec or approximately 1.08 h. If we wish to consider three stars of masses $3M_\odot$, $4M_\odot$, and $5M_\odot$ at distances 3, 4, and 5 pc, then the unit of time becomes about 1.43×10^7 yr. The unit of length $L = 1$ pc = 3×10^{18} cm in this last example corresponds to one third of the average distance (3 pc) between the stars in the neighborhood of the sun ($M_\odot = 2 \times 10^{33}$ g).

The initial motion of the bodies is expected to be oriented toward the center of mass. At $t_1 = 1.879$ this trend ends, and as shown in Fig. 2 the second and third bodies approach each other, while the first body continues its orbit. The closest distance between m_2 and m_3 is approximately $r_{23}(t_1) \cong 10^{-2}$. Following the orbits for $0 \leq t \leq 10$ of the three bodies in Fig. 2, the observation may be made that the third body with the largest mass performs repeated close approaches to the first and second bodies but these latter bodies do not find themselves near to each other. This tendency of the body with the largest mass to act as an agent between the two other masses persists during the entire motion. In other words, the communication and the direct gravitational effect between m_1 and m_2 is only through m_3 . This phenomenon of separation between m_1 and m_2 may be seen clearly in Figs. 2, 3, and 4. In Fig. 7 there is an instant when the two bodies with

FIG. 3. Orbits between $t=10$ and $t=20$.FIG. 4. Orbits between $t=20$ and $t=30$.

the smaller masses are closer to each other than to the third body ($r_{12} < r_{23}$). As discussed later, m_2 and m_3 at this critical time are already in the process of forming a close binary system that will be their final configuration.

Returning now to Fig. 2, we direct attention to the system at $t=3.35$. At this time m_1 moves away from the origin, while m_2 and m_3 are on approach trajectories. Since Burrau's integration ended at $t=3.35$, it is understandable that he could not find any indication of either periodicity or of an asymptotic behavior.

The close approaches between m_2 and m_3 in Fig. 2 occur in the following intervals of time: $1 < t_1 < 2$, $3 < t_2 < 4$, and $8 < t_3 < 9$. The corresponding distances are $r_{23}(t_1) \cong 10^{-2}$, $r_{23}(t_2) \cong 6 \times 10^{-2}$, and $r_{23}(t_3) \cong 8 \times 10^{-3}$. The last close approach occurs at $t_3=8.760$ and it is the

smallest distance during the interval of time shown in this figure. The close approaches between m_1 and m_3 occur at $t_4 \cong 3$, $6 < t_5 < 7$, and at $t_6 \cong 10$. The distances at the close approaches between m_1 and m_3 are all larger than the close approaches between m_2 and m_3 for $0 \leq t \leq 10$.

Proceeding now to Fig. 3, a clear separation of the orbits of m_1 and m_2 , with m_3 oscillating between them, may be observed. The close approaches between m_1 and m_3 occur at $t_6 \cong 10$, $14 < t_7 < 15$, and at $t_8 \cong 17$. The corresponding times for m_2 and m_3 are $11 < t_9 < 12$, $t_{10}=15.8299236$, and $19 < t_{11} < 20$. The close approach at t_{10} is $r_{23}(t_{10}) \cong 4 \times 10^{-4}$ which is the smallest distance occurring between any bodies at any time during the evolution of this dynamic system between $t=0$ and

$t = \infty$. The relative velocity of m_2 and m_3 at close approach is approximately 171 in our units and at the same time the velocity of m_1 is $v_1 = 0.04$.

Note that this closest approach corresponds to a distance of 1.7×10^4 solar radius or 80 a.u. and the maximum velocity becomes approximately 10 km/sec. The nearly stationary condition of m_1 is reflected by its velocity of approximately 10 km/h.

If the distance r_{23} would have become zero and at the same time v_1 would also have obtained the value of zero, a periodic solution would have presented itself. In other words, had m_2 and m_3 collided and m_1 stopped at the same time, the solution would have reversed. Such a tendency may be seen in Fig. 3 where all three bodies describe approximately the same orbits before and after the close approach. This trend continues through Figs. 4 and 5 since half of the period would be $t_{10} = 15.83 = \frac{1}{2}T$ if the motion would be periodic. Indeed, in Fig. 5 at $t_{12} = 2t_{10} = 31.66 = T$, the three bodies approximately occupy their initial positions with small velocities. The existence of a periodic solution near the Pythagorean initial conditions used in this paper is still being investigated. Note that the period of the three-body problem mentioned corresponds approximately to twice the period of rotation of the galaxy.

Figure 4 resembles part of Fig. 3 and part of Fig. 2 because of the near periodicity described in the preceding paragraph. Figure 5 at $t_{12} = 31.66$ resembles Fig. 2 at $t=0$. The sensitivity of this dynamical system to initial conditions is well demonstrated in Fig. 5. Since the initial conditions of Fig. 2 are not exactly reproduced in Fig. 5 at t_{12} , the subsequent motion in these two figures is entirely different. In other words, no similarity may be seen between Figs. 2 and 5 after t_{12} .

The next figure already foreshadows the final evolution of the system. At $t_{13} = 47$ in Fig. 6 the body with the smallest mass m_1 seems to be departing in the third quadrant of the x, y plane with $\dot{x}_1 < 0, \dot{y}_1 < 0$. At the same time m_2 and m_3 form a binary which is moving away from the center of mass in the opposite direction. That ejection does not occur at this time becomes clear when Fig. 7 is inspected. At $t_{14} = 53$ the first body reverses its direction and since its velocity is almost zero, it evidently has an elliptic-type motion with respect to the binary.

Figure 7 shows the final formation of a binary by m_2 and m_3 and the penetration of this binary by m_1 . At $t_{15} = 54$, the body with the smallest mass m_1 approaches the binary with $\dot{x}_1 > 0, \dot{y}_1 > 0$. At $t_{16} = 59.2$ the binary is in the position of periastron and at $t_{17} = 59.4$ the condition is almost that of apastron. At the same time m_1 is approximately on the line connecting m_2 and m_3 . After this, m_1 departs fast while m_2 and m_3 reach periastron again, approximately at $t_{18} = 59.7$. One way to describe the interaction may be by referring to m_1 as shooting through the binary without influencing it very much and attaining a high velocity in the process. At $t_{19} = 60$ the speed of m_1 is approximately 2.7 and its

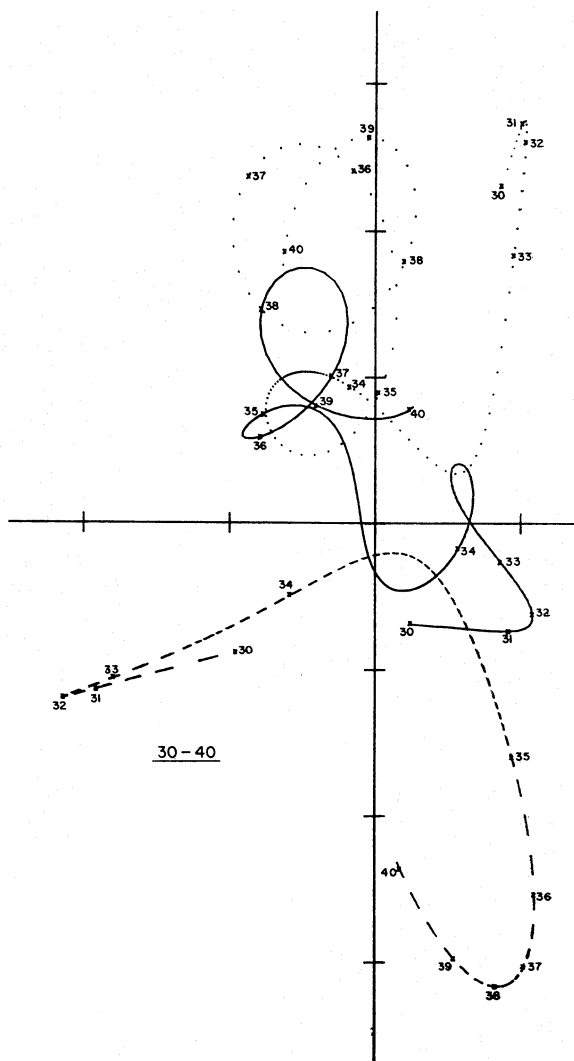


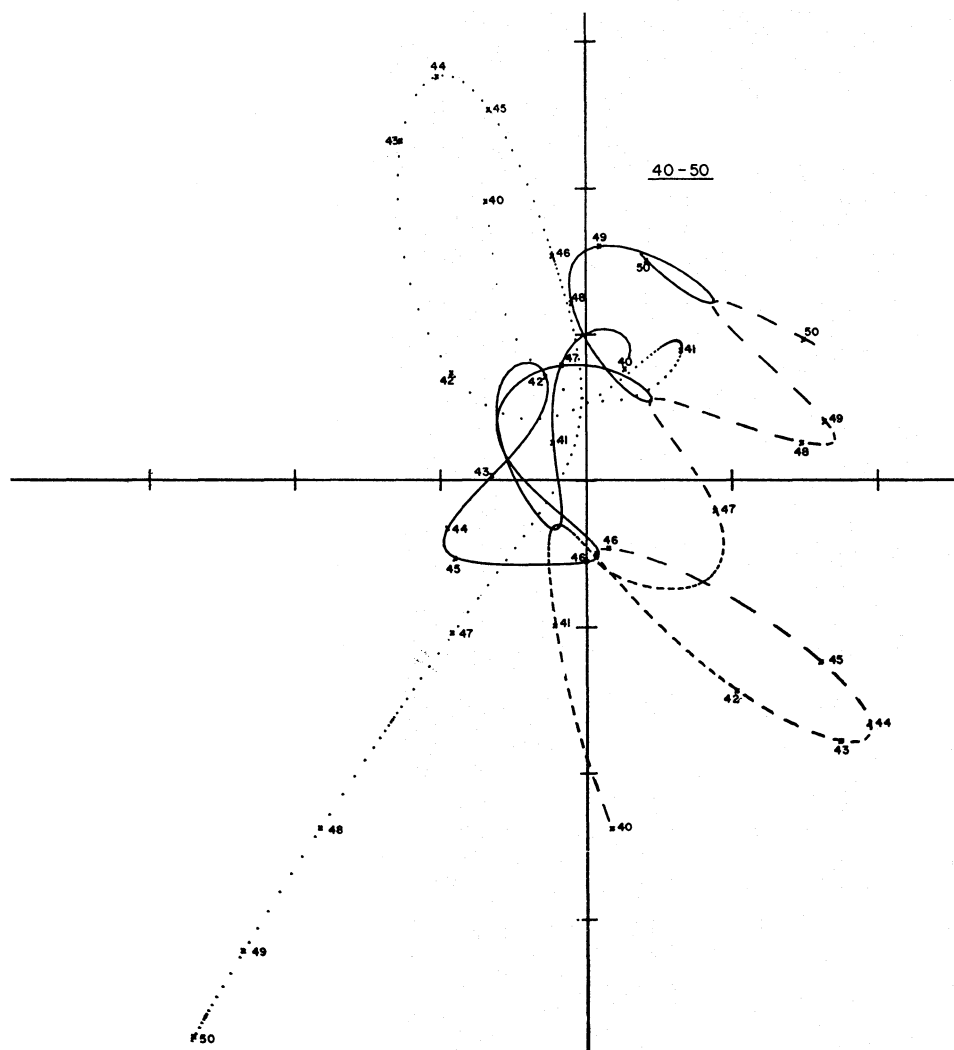
FIG. 5. Orbits between $t=30$ and $t=40$.

distance from the center of mass of the system is about 2 units. If one considers a problem of two bodies, the first body being m_1 and the second the binary with mass $m_2 + m_3$ and with velocity 0.9, the total energy of this artificial system becomes positive; consequently the motion is hyperbolic.

Figure 8 completes the problem. The binary and the body with mass m_1 depart with hyperbolic velocities. The period of the binary is approximately 0.9 time unit and therefore this binary performs 15 revolutions during one complete rotation of the galaxy.

The previously mentioned approximate calculation regarding the hyperbolic departure may be sharpened by using Alexeev's (1961) conditions. In this way we find that his conditions of hyperbolic-elliptic motion are satisfied at $t=69$. The numerical integration was performed to $t_{20} = 102$ or to 1.5×10^9 yr and gave no change from the situation shown in Fig. 8.

Table I lists the data for the close approaches

FIG. 6. Orbits between $t=40$ and $t=50$.

occurring between $t=0$ and $t=30$, i.e., for the first 430 million years.

METHOD OF REGULARIZATION

From preliminary studies, it was found that several very close approaches occurred during the course of the numerical integration. In order to increase accuracy and to save time spent on a high-speed electronic computer, the equations of motion were subjected to the treatment known as regularization, which seems to be the only available method to deal with collisions and close approaches. The basic idea is to eliminate the singularities occurring at binary collisions in the problem of three bodies under consideration by introducing a new independent variable. This method has been proposed by Sundman (1912) and the effectiveness of the transformation is unquestionable from the point of view of analysis. Unfortunately, for purposes of

numerical integration Sundman's method and other similar approaches do not seem to be satisfactory because of practical computational reasons. One primary reason for this discrepancy between the analytic and computational requirements is the appearance of terms of the type x_i/r in the transformed equations. Here $r = [\sum(x_i^2)]^{1/2}$ and x_i and r both approach zero at collision. The limit of the ratio x_i/r is well defined and entirely satisfactory for analytical purposes, nevertheless its time-consuming computation gives losses in accuracy. This disadvantage of Sundman's transformation is well known from the numerical experiences of other dynamical systems, such as the restricted problem of three bodies. It has been shown, for instance by one of us (Szebehely 1967), that the introduction of a new independent variable, while regularizing the restricted problem, results in increased complexity of the equations of motion as well as being responsible for the appearance of terms undesirable from the point

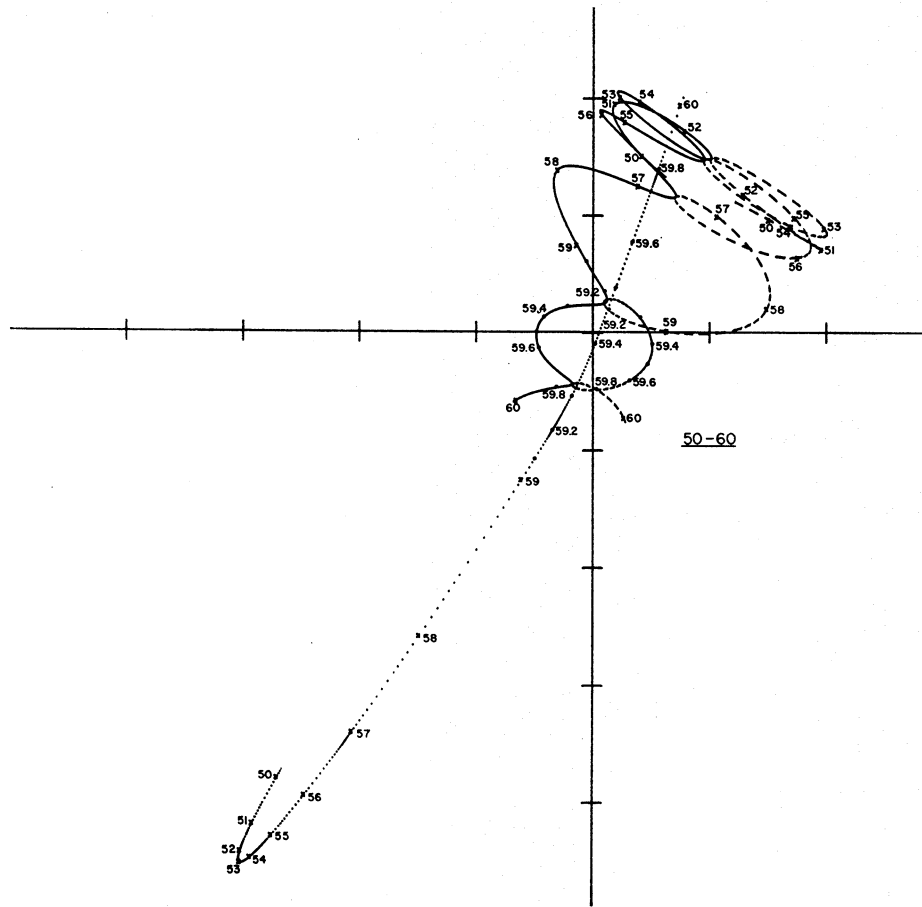


FIG. 7. Orbits between $t=50$ and $t=60$.

of view of numerical integration. The remedy is to increase the complexity of the regularizing transformations in the case of the restricted problem, where the proper combination of the transformations of the independent and dependent variables accomplishes not

only regularization but also significant simplifications of the equations of motion. The generalization of this idea to the gravitational problem of n bodies will be expounded elsewhere. It should be mentioned that the need for the transformation of the dependent and independent variables was conjectured by Lemaitre (1955) by stating "It is possible that it is the introduction of the superfluous conformal transformation which makes the whole difference between a mere mathematical theorem and a powerful numerical tool."

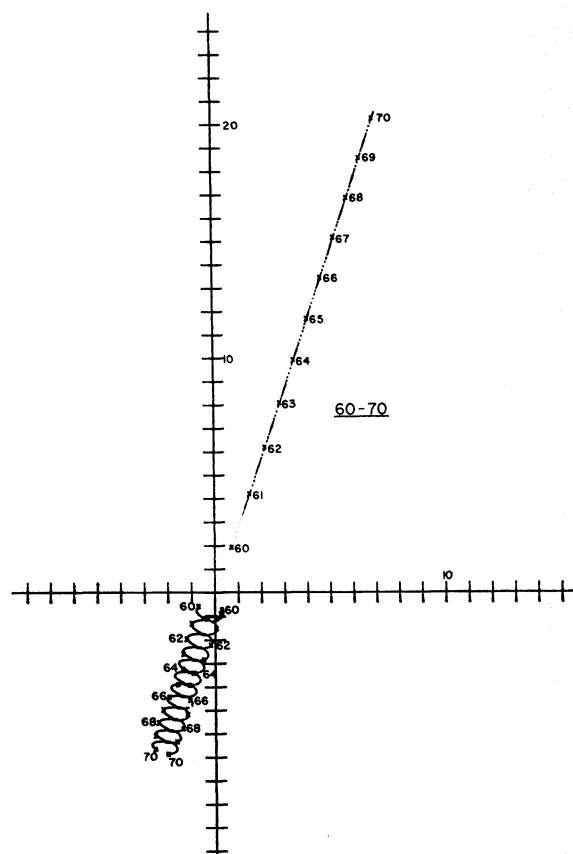
TABLE I. Close approaches between $t=0$ and $t=30$.

time	approximate distances	between particles
$t_1 = 1.879$	10^{-2}	m_2, m_3
$t_4 = 3.026$	0.6	m_1, m_3
$t_2 = 3.801$	6×10^{-2}	m_2, m_3
$t_5 = 6.898$	0.1	m_1, m_3
$t_3 = 8.760$	8×10^{-3}	m_2, m_3
$t_6 = 9.962$	0.5	m_1, m_3
$t_9 = 11.611$	0.2	m_2, m_3
$t_7 = 14.618$	0.2	m_1, m_3
$t_{10} = 15.830$	4×10^{-4}	m_2, m_3
$t_8 = 17.001$	0.3	m_1, m_3
$t_{11} = 19.807$	0.2	m_2, m_3
21.791	0.4	m_1, m_3
22.966	2×10^{-2}	m_2, m_3
24.537	0.1	m_1, m_3
27.780	5×10^{-2}	m_2, m_3
28.679	0.5	m_1, m_3
29.802	3×10^{-3}	m_2, m_3

A special method is mentioned here which seems to be well suited for the solution of the planar problem of three bodies. The transformation used was originally proposed by Levi-Civita (1904), and it combines the introduction of a new time-variable τ^* and the transformation of the coordinates from the system x, y to the system ξ, η . The essential aspects are described by the equations

$$\tau^* = \int_{t_0}^t \frac{dt}{r} + \tau_0^* \quad \text{and} \quad x + iy = (\xi + i\eta)^2.$$

The application of these equations to the general planar problem of three bodies is not at all trivial. It

FIG. 8. Orbits between $t=60$ and $t=70$.

is first necessary to transform the equations of motion into a form suitable for the application of Levi-Civita's method. A convenient choice is a system in which x and y are the relative coordinates of one colliding body with respect to the other. The method is used for the regularization of the restricted problem and the reader may be referred to standard references such as for instance to Wintner (1941) in order to see how the method is applied to a relatively simple problem.

The following few paragraphs will indicate the application of Levi-Civita's method to the actual numerical integration of our problem. Upon deciding that only isolated binary collisions are to be regularized, a limiting value for r_{ij} is set, below which the transformation is applied. Here r_{ij} is the distance between the i th and j th particles and, therefore, the previously mentioned limit means that when two of the three bodies are closer than this limiting value, the equations of motion of these two bodies are to be regularized. If all three distances are below the limit, that pair is selected for regularization which has the smallest distance. At the time when the distance between a critical pair decreases below the given limit, the equations are regularized and when the pair increases their distance above the limit, the equations of motion are transferred back to the original variables. This pro-

cedure, therefore, requires that regularization be switched on and off quite often, the frequency depending on the limit set and on the special dynamical system under consideration. Attention is called to the complications introduced by changing the variables during the course of the numerical integration and to the fact that every time such a change is made, the integration must be restarted with a new set of initial conditions. Runge-Kutta's method of integration is well adaptable to this special requirement.

METHOD OF NUMERICAL INTEGRATION

A fifth-order Runge-Kutta scheme developed by Zonneveld (1964) was used to numerically integrate the equations of motion. Zonneveld also gave an expression for the last term, of the fifth order, included in the Taylor series which was used to derive the algorithm. This term may be used as an estimate of the truncation error. Ollongren (1966) has derived an elegant formula for predicting the optimum step size to be used at each step of the integration. If h_i is the size of the i th step, E_{\max} the maximum allowable truncation error, and E_5 the estimate of the truncation error of the fifth-order scheme, then the optimum size of the $(i+1)$ th step is given by

$$h_{i+1} = \left(\frac{E_{\max}}{E_{\max} + E_5} + 0.45 \right) h_i. \quad (2)$$

In the case where $E_5 \geq E_{\max}$ the step is rejected, and the formula is used again until the requirement for the tolerance is satisfied, i.e., $E_5 < E_{\max}$. In the problem discussed, about one out of every 1000 steps was rejected.

The value of the total energy was monitored throughout the integration. In addition, the conservation of the angular momentum and the integral of the center of mass were also checked at selected times. In all cases the requirement of keeping the total energy constant proved to be the most sensitive control. It should be remarked, however, that the check of the integral of energy is not always reliable, being a necessary but not a sufficient condition for accuracy.

In the regularized system the total energy is defined in terms of the original variables and, as such, it is an inconvenient quantity to compute. In this case the transformed Hamiltonian, which should be identically zero, is used as the control of the energy.

COMPARISON AND ACCURACY

Some aspects of the accuracy of the numerical integration have been discussed in the preceding section. Here the question of reversibility is taken up since its numerical and dynamic aspects are not always clearly separable. The dynamical system under consideration is reversible. The equations contain the independent

variable t in such a manner that a change from t to $(-t)$ does not change the equations of motion. Consequently, it is expected that when the differential equations representing the dynamic system are integrated from, say, t_0 to t_1 and at t_1 integration is reversed to t_0 , the original initial conditions are obtained. (The integrations in the forward and reverse directions are completely independent, since the integration interval is computed at each time step. Only by pure chance would the times in the two integrations agree for any particular time.) In this process no numerical errors are allowed for, since these inaccuracies may negate the reversibility of the system. The process of regularization is helpful in this respect since it allows a significant increase in the accuracy. Consequently, when actual reversibility is considered to be the measure of the accuracy of a numerical integration, efficient regularization becomes mandatory. In the problem treated in this paper reversals were executed at $t=32$ and at $t=62$. The initial conditions for the positions were reconstructed for these two cases with errors in the tenth and third decimals, respectively, keeping the error in the total energy below 10^{-10} in both cases. Reversal at $t=62$ required about 19 000 steps of integration.

It is of interest to compare various numerical methods of solution for the Pythagorean problem treated in this paper and therefore the speed and accuracy of the numerical integrations for three different systems are shown in the range $0 \leq t \leq 25$, where 14 close approaches occur. In the first system, denoted by R , the conventional system of rectangular coordinates is used as introduced before. In the system denoted by V the potential energy is used as the regularizing function and only the time is regularized. The system denoted by L corresponds to the variables introduced by Levi-Civita.

Table II presents the results of four computations. The numbers listed under the heading of tolerance represent the accuracy attained in the first few integration steps. In systems R and V the accuracy degenerated at each close encounter and at a much faster rate than for system L . This process of degeneration is a function of how close the bodies approach one

TABLE II. Comparison of different systems.

System	Number of steps	Time spent on computer (sec)	Change in total energy	Tolerance
R	17 631	395	7×10^{-7}	10^{-12}
V	4300	280	8×10^{-7}	10^{-12}
L	3450	155	2×10^{-11}	10^{-10}
L	600	26	7×10^{-8}	10^{-6}

another when systems R or V are used since all numerical accuracy would be lost if a collision should occur. In system L , however, it is possible to integrate through collisions without difficulty. It was found that greater computer speed could be achieved if the L system was used when the particles were less than a distance of 2 apart. From the figures showing the trajectories this means that the integration was in the L system during the majority of the time.

ACKNOWLEDGMENTS

This research was partially sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research U. S. Air Force under AFOSR Grant No. 397-67 and by the Office of Naval Research, U. S. Navy under Contract No. ONR 609(50). The second author wishes to acknowledge the sponsorship of the U. S. Naval Weapons Laboratory, Dahlgren, Virginia.

REFERENCES

- Alexeev, V. M. 1961, *Astron. J. U.S.S.R.* 38, 1099; also 1962, *Soviet Astron.* 5, 841.
 Burrau, C. 1913, *Astron. Nachr.* 195, 113.
 Lemaitre, G. 1955, *Vistas in Astronomy*, A. Beer, Ed., Vol. 1, p. 207.
 Levi-Civita, T. 1904, *Ann. Mat. Ser. 3*, 9, 1.
 Merman, G. A. 1958, *Bull. Inst. Theoret. Astron. Leningrad* 6, 687.
 Ollongren, A. 1966 (private communication).
 Spinelli, R. 1966 (private communication).
 Stanek, L. 1966 (private communication).
 Standish, E. M. 1966 (private communication).
 Sundman, C. F. 1912, *Acta Math.* 36, 105.
 Szebehely, V. 1967, *Astron. J.* 72, 370.
 ——. 1967, *Proc. Natl. Acad. Sci. U. S.* 57, July.
 Wintner, A. 1941, *The Analytical Foundation of Celestial Mechanics* (Princeton University Press, Princeton, New Jersey).
 Zonneveld, A. 1964, *Mathematical Tracts No. 8*, Mathematical Center, Amsterdam.