Lectures On
The Singularities Of The
Three–Body Problem

By

C.L. Siegel

Tata Institute of Fundamental Research, Bombay
1967

Lecture on
The singularities of
the three–body problem

By

C. L. Siegel

Notes by

K. Balagangadharan

and

M. K. Venkatesha Murthy

Tata Institute of Fundamental Research, Bombay
1967

Contents

1 The differential equations of mechanics **1**
- 1 The Euler-Lagrange equations . . . 1
- 2 The transformation theory of Hamiltonian equations . . . 11
- 3 Cauchy's theorem on the existence of solutions... . . . 21

2 The three-body problem : simple collisions **31**
- 1 The n-body problem . . . 31
- 2 Collisions . . . 39
- 3 Simple collisions in the case $n = 3$ . . . 45
- 4 Reduction of the differential equations of motion . . . 52
- 5 Approximate solution of the Hamilton-Jacobi... . . . 57
- 6 Regularisation of the solution of the three-body... . . . 66

3 The three-body problem: general collision **75**
- 1 Asymptotic estimates . . . 75
- 2 The limiting configuration at a general collision . . . 87
- 3 A particular solution . . . 100
- 4 Reduction to a rotating coordinate system . . . 104
- 5 Stability theory of solutions of differential equations . . . 126
- 6 Application to the three-body problem . . . 175

Chapter 1

The differential equations of mechanics

1 The Euler-Lagrange equations

We shall begin with some introductory remarks on the differential equations of mechanics; we shall indicate their connection with the calculus of variatinos and discuss briefly the transformatin theory due to Hamilton and Jacobi. 1

Let n be a positive integer and let $x = (x_1, \ldots, x_n)$, $\dot{x} = (\dot{x}_1, \ldots, \dot{x}_n)$ and t be $2n + 1$ independent real variables. Suppose that $f(x, \dot{x}, t)$ is a real twice continuously differentiable function of the $2n + 1$ variables $(x, \dot{x}, t)$ for $t_1 \leq t \leq t_2$ and $(x, \dot{x})$ belonging to an open set G of $2n$-dimensional Euclidean space. Next, let us suppose that each x_k, $k = 1, \ldots, n$, is a twice continuously differentiable real-valued function $x_k(t)$ of the variable t in $t_1 \leq t \leq t_2$; we set $\dot{x}_k(t) = \dfrac{dx_k(t)}{dt}$. We also write $x(t) = (x_1(t), \ldots, x_n(t))$, $\dot{x}(t) = (\dot{x}_1(t), \ldots, \dot{x}_n(t))$, and assume that $(x(t), \dot{x}(t)) \in G$ for $t_1 \leq t \leq t_2$. Then $f(x(t), \dot{x}(t), t)$ is a continuous (in fact, twice continuously differentiable) function of the variable t in $t_1 \leq t \leq t_2$ and so we can form the integral

$$\int_{t_1}^{t_2} f(x(t), \dot{x}(t), t)dt. \tag{1.1.1}$$

The classical problem of the calculus of variations consists in determin-
2 ing a twice continuously differentiable function $x = x(t)$ of the variable t in $t_1 \leq t \leq t_2$ with $(x(t), \dot{x}(t)) \in G$ and satisfying prescribed initial conditions $x(t_1) = a$ and $x(t_2) = b$, where a and b are given points in n-dimensional Euclidean space, such that the integral (1.1.1) is a minimum. Let us assume that this minimising problem has a solution. We derive a necessary condition for the existence of such a solution $\overline{x} = \overline{x}(t)$. Let $y = y(t) = (y_1(t), \ldots, y_n(t))$ be a fixed twice continuously differentiable function of the variable t in $t_1 \leq t \leq t_2$ with $y(t_1) = 0$ and $y(t_2) = 0$. Since G is open, if $x = x(t, \epsilon) = \overline{x}(t) + \epsilon y(t)$, then $(x, \dot{x}) \in G$ for all real ϵ with sufficiently small absolute value. Moreover, we have $x(t_1, \epsilon) = a$ and $x(t_2, \epsilon) = b$ for all ϵ. Then the integral

$$\int_{t_1}^{t_2} f(x(t, \epsilon), \dot{x}(t, \epsilon), t)\ dt,$$

where $\dot{x}(t, \epsilon)$ denotes $\dfrac{d}{dt}x(t, \epsilon)$, defines a real-valued function $J = J(\epsilon)$ of the parameter ϵ. Since $f(x, \dot{x}, t)$ is twice continuously differentiable in all the $2n + 1$ variables $(x, \dot{x}, t)$ and $x(t, \epsilon)$ is linear in ϵ, the function $J(\epsilon)$ is a twice continuously differentiable function of ϵ. Further, since $\overline{x} = x(t, 0)$ is a minimising function for the integral (1.1.1), $J(0)$ is a minimum value of the function $J(\epsilon)$ A necessary condition that $J(\epsilon)$ has a minimum value at $\epsilon = 0$ is that $\dfrac{dJ(\epsilon)}{d\epsilon} = 0$ at $\epsilon = 0$. From this one gets the Euler-Lagrange equations in the following way. We have

$$\frac{dJ(\epsilon)}{d\epsilon} = \int_{t_1}^{t_2} \frac{\partial}{\partial\epsilon} f(x(t, \epsilon), \dot{x}(t, \epsilon), t)dt,$$

3 and if, for any fixed value of the parameter ϵ, f_{x_k} and $f_{\dot{x}_k}$ denote the

partial derivatives $\dfrac{\partial f(x,\dot{x},t)}{\partial x_k}$ and $\dfrac{\partial f(x,\dot{x},t)}{\partial \dot{x}_k}$ of the function f, considered as a function of the $2n+1$ variables $(x,\dot{x},t)$, then we have, by the chain-rule for differentiation,

$$\begin{aligned}\frac{\partial}{\partial \epsilon} f(x,\dot{x},t) &= \sum_{k=1}^{n}\Big(f_{x_k}\frac{\partial x_k(t,\epsilon)}{\partial \epsilon} + f_{\dot{x}_k}\frac{\partial \dot{x}_k(t,\epsilon)}{\partial \epsilon}\Big)\\ &= \sum_{k=1}^{n}\big(f_{x_k}(x,\dot{x},t)y_k(t) + f_{\dot{x}_k}(x,\dot{x},t)\dot{y}_k(t)\big).\end{aligned}$$

Hence we obtain

$$\frac{dJ(\epsilon)}{d\epsilon} = \int_{t_1}^{t_2}\sum_{k=1}^{n}\big(f_{x_k}(x,\dot{x},t)y_k(t) + f_{\dot{x}_k}(x,\dot{x},t)\dot{y}_k(t)\big)dt,$$

and this gives, on integration by parts,

$$\begin{aligned}\frac{dJ(\epsilon)}{d\epsilon} = \int_{t_1}^{t_2}\sum_{k=1}^{n}\Big(f_{x_k}(x,\dot{x},t) - \frac{d}{dt}f_{\dot{x}_k}(x,\dot{x},t)\Big)y_k(t)dt+\\ + \int_{t_1}^{t_2}\frac{d}{dt}\left(\sum_{k=1}^{n} f_{\dot{x}_k}(x,\dot{x},t)y_k(t)\right)dt.\end{aligned}$$

The second term on the right side vanishes since $y_k(t_1) = 0 = y_k(t_2)$. Therefore a necessary condition for $\dfrac{dJ(\epsilon)}{d\epsilon}$ to vanish at $\epsilon = 0$ is that all **4**
the $f_{x_k} - \dfrac{d}{dt}f_{\dot{x}_k}$, $k = 1,\ldots,n$, vanish for $x = x(t,0) = \overline{x}(t)$. Thus we obtain a system of differential equations

$$\wedge_k(f) \equiv f_{x_k} - \frac{d}{dt}f_{\dot{x}_k} = 0,\ k = 1,\ldots,n; \qquad (1.1.2)$$

these are the *Euler-Lagrange differential equations.*

We can rewrite the system (1.1.2) of differential equations more explicitly using the fact that the function f is twice continuously differentiable in all its $2n+1$ independent variables $(x,\dot{x},t)$. Carrying out the differentiation with respect to t we get

$$\wedge_k(f) = f_{x_k} - \sum_{l=1}^{n}(f_{\dot{x}_k x_1}\dot{x}_1 + f_{\dot{x}_k\dot{x}_1}\ddot{x}_1) - f_{\dot{x}_k t} = 0,$$

where

$$\ddot{x}_k = \frac{d^2 x_k(t)}{dt^2},\ f_{\dot{x}_k t} = \frac{\partial^2 f(x,\dot{x},t)}{\partial\dot{x}_k\partial t},\ f_{\dot{x}_k\dot{x}_1} = \frac{\partial^2 f(x,\dot{x},t)}{\partial\dot{x}_k\partial\dot{x}_1},\ f_{\dot{x}_k x_1} = \frac{\partial^2 f(x,\dot{x},t)}{\partial\dot{x}_k\partial x_1},$$

$k, l = 1, \ldots, n$. This is a system of n partial differential equations of the second order in n unknown function $x_k(t)$, $k = 1, \ldots, n$.

If $f = f(x, \dot{x}, t)$ is an arbitrary twice continuously differentiable function of $2n + 1$ independent real variables $(x, \dot{x}, t)$ in $t_1 \leq t \leq t_2$ and $(x, \dot{x}) \in G$, and if $x = x(t)$ is a twice continuously differentiable function of t in $t_1 \leq t \leq t_2$ such that $(x(t), \dot{x}(t)) \in G$ (where $\dot{x}_k(t) = \frac{d}{dt}x_k(t)$, $k = 1, \ldots, n$), then we can form the expression

$$\wedge_k(f) \equiv f_{x_k} - \sum_{l=1}^{n}(f_{\dot{x}_k x_1}\dot{x}_1 + f_{\dot{x}_k\dot{x}_1}\ddot{x}_1) - f_{\dot{x}_k t}, k = 1, \ldots, n;$$

this is called the *Lagrangian derivative* of f.

We can try to simplify the Euler-Lagrange equations by means of a suitably chosen substitution. We introduce new independent variables $\xi = (\xi_1, \ldots, \xi_n)$ and set

$$x_k = x_k(\xi, t),\ k = 1, \ldots, n, \tag{1.1.3}$$

where $x_k(\xi, t)$ are twice continuously differentiable functions of the $n+1$ independent real variables (ξ, t), and we assume that the Jacobian of x with respect to ξ does not vanish anywhere in the region under consideration, so that the transformation from ξ to x is locally one-to-one everywhere. Thus, writing $x_{k\xi_1} = \frac{\partial x_k}{\partial \xi_1}$, we assume that the determinant $|x_{k\xi_1}| \neq 0$ everywhere. Then we can invert the transformation (1.1.3) locally and determine ξ_1 as a function $\xi_1(x, t)$ of the $n + 1$ independent variables (x, t), $l = 1, \ldots, n$. We can write, by the chain-rule for differentiation,

$$\dot{x}_k = \sum_{r=1}^{n} x_{k\xi_r}\dot{\xi}_r + x_{kt}, \tag{1.1.4}$$

$$\ddot{x}_k = \sum_{r=1}^{r}\left(\sum_{l=1}^{n} x_{k\xi_r\xi_1}\dot{\xi}_1\dot{\xi}_r + x_{k\xi_r t}\dot{\xi}_r + x_{k\xi_r}\ddot{\xi}_r\right) + x_{ktt}, \tag{1.1.5}$$

where

$$\dot{\xi}_r = \frac{d}{dt}\xi_r(t), \ddot{\xi}_r = \frac{d^2}{dt^2}\xi_r(t), x_{ktt} = \frac{\partial^2 x_k(\xi, t)}{\partial t^2},$$
$$x_{k\xi_1\xi_r} = \frac{\partial^2 x_k(\xi, t)}{\partial\xi_1\partial\xi_r} \quad \text{and} \quad x_{k\xi_r t} = \frac{\partial^2 x_k(\xi, t)}{\partial\xi_r\partial t}.$$

(We may also consider $\dot{\xi}_1, \ddot{\xi}_1$ as independent variables and define $\dot{x}_k$, $\ddot{x}_k$ **6**
in terms of $\dot{\xi}_1, \ddot{\xi}_1$ by means of the equations (1.1.4) and (1.1.5)).

We suppose that $(\xi, \dot{\xi})$ varies in an open set G_1 in $2n$-dimensional Euclidean space so that $(x, \dot{x}) \in G$. Then we define the function $g = g(\xi, \dot{\xi}, t)$ by

$$g(\xi, \dot{\xi}, t) = f(x(\xi, t), \dot{x}(\xi, \dot{\xi}, t), t).$$

The function g is again twice continuously differentiable in the variables $(\xi, \dot{\xi}, t)$. We shall now obtain the relation between the Lagrangian derivatives $\wedge_k(f)$ and $\wedge_k(g)$ of f and g. We have, by definition,

$$\wedge_1(g) = g_{\xi_1} - \frac{d}{dt}g_{\dot{\xi}_1},$$

and by the chain-rule for differentiation, using (1.1.4), we get

$$g_{\xi_1} = \sum_{k=1}^{n}\left(f_{x_k}x_{k\xi_1} + f_{\dot{x}_k}\left(\sum_{r=1}^{n} x_{k\xi_r\xi_l}\dot{\xi}_r + x_{kt\xi_1}\right)\right),$$
$$g_{\dot{\xi}_1} = \sum_{k=1}^{n} f_{\dot{x}_k}(\dot{x}_k)\dot{\xi}_1 = \sum_{k=1}^{n} f_{\dot{x}_k}x_{k\xi_1},$$
$$\frac{d}{dt}g_{\dot{\xi}_1} = \sum_{k=1}^{n}\left(\frac{d}{dt}f_{\dot{x}_k}x_{k\xi_1} + f_{\dot{x}_k}\sum_{r=1}^{n}(x_{k\xi_1\xi_r}\dot{\xi}_r + x_{k\xi_1 t})\right)$$

Since $x_k = x_k(\xi, t)$ is twice continuously differentiable in (ξ, t),

$$x_{kt\xi_1} = x_{k\xi_1 t} \quad \text{and} \quad x_{k\xi_r\xi_1} = x_{k\xi_1\xi_r},$$

7 and consequently,

$$\wedge_1(g) = \sum_{k=1}^{n}(f_{x_k} - \frac{d}{dt} f_{\dot{x}_k})x_{k\xi_1} = \sum_{k=1}^{n} \wedge_k(f)x_{k\xi_1}.$$

As a result, if $\wedge_k(f) = 0$, $k = 1, \ldots, n$, and if $x_k = x_k(\xi, t)$ are twice continuously differentiable functions of (ξ, t) with the determinant $|x_{k\xi_1}| \neq 0$, then the function $g(\xi, \dot{\xi}, t) = f(x, \dot{x}, t)$ satisfies the relation $\wedge_1(g) = 0$, $l = 1, \ldots, n$, and conversely. This relation could also be obtained directly by considering the integral

$$\int_{t_1}^{t_2} g(\xi, \dot{\xi}, t)dt = \int_{t_1}^{t_2} f(x, \dot{x}, t)dt$$

and differentiating partially.

We can write the relation obtained above:

$$\wedge_1(g) = \sum_{k=1}^{n} \wedge_k(f)x_{k\xi_1}, \tag{1.1.6}$$

in a simpler way as follows. If we denote the Jacobian matrix $(x_{k\xi_1})$ by M and write

$$\wedge(f) = (\wedge_1(f), \ldots, \wedge_n(f)); \;\; \wedge(g) = (\wedge_1(g), \ldots, \wedge_n(g)),$$

then

$$\wedge(g) = \wedge(f)M.$$

This is the covariance property of the Lagrangian derivative.

8 The Lagrangian derivative $\wedge_k(f)$ of a function f can be considered as a function of $3n + 1$ independent real variables $(x, \dot{x}, \ddot{x}, t)$. Now we investigate the condition under which the relation $\wedge_k(f) = \wedge_k(h)$ holds identically in all the $3n + 1$, independent variables for two functions

$f(x, \dot{x}, t)$ and $h(x, \dot{x}, t)$ which are twice continuously differentiable in the $2n + 1$ variables $(x, \dot{x}, t)$ for $t_1 \le t \le t_2$ and $(x, \dot{x}) \in G$.

Since the Lagrangian derivatives are linear homogeneous operators, it is enough, setting $s = h - f$, to investigate when $\wedge_k(s) = 0$, $k = 1, \ldots, n$, identically in the $3n + 1$ independent variables $x, \dot{x}, \ddot{x}, t$. Since

$$\wedge_k(s) = s_{x_k} - \sum_{l=1}^{n} (s_{\dot{x}_k x_l} \dot{x}_1 + s_{\dot{x}_k \dot{x}_1} \ddot{x}_1) - s_{\dot{x}_k t},$$

it follows that the coefficient of $\ddot{x}_1$ vanishes identically, i.e.

$$s_{\dot{x}_k \dot{x}_1} \equiv 0, \quad k, l = 1, \ldots, n,$$

which means that s is a linear function of $(\dot{x}_1, \ldots, \dot{x}_n)$. Hence s has the form

$$s(x, \dot{x}, t) = \sigma_\circ(x, t) + \sum_{l=1}^{n} \sigma_1(x, t) \dot{x}_1,$$

where $\sigma_\circ(x, t)$, $\sigma_1(x, t)$ are twice continuously differentiable functions of (x, t). So $\wedge_k(s) \equiv 0$ gives

$$\sigma_{\circ x_k} + \sum_{l=1}^{n} \sigma_{l x_k} \dot{x}_l - \sum_{l=1}^{n} \sigma_{k x_1} \dot{x}_l - \sigma_{kt} \equiv 0.$$

This implies that **9**

$$\sigma_{\circ x_k} = \sigma_{kt} \quad \text{and} \quad \sigma_{1 x_k} = \sigma_{k x_1}, k, l = 1, \ldots, n.$$

If we define for the moment $x_\circ = t$, then the first condition becomes $\sigma_{\circ x_k} = \sigma_{k x_\circ}$, $k = 1, \ldots, n$, which is of the same form as the other conditions. These are necessary and sufficient conditions in order that there exist a function $\sigma(x, t)$, twice continuously differentiable in the $n + 1$ independent variables $(x_\circ, x_1, \ldots x_n)$, such that

$$\sigma_k(x, x_\circ) = \frac{\partial \sigma(x, x_\circ)}{\partial x_k}, \quad k = 0,\ 1, \ldots n.$$

Hence, under these condition there exists a function $\sigma(x,t)$, twice continuously differentiable in the $n+1$ independent variables (x,t), such that

$$\sigma_\circ(x,t) = \frac{\partial\sigma(x,t)}{\partial t} \quad \text{and} \quad \sigma_k(x,t) = \frac{\partial\sigma(x,t)}{\partial x_k}, k = 1,\ldots,n.$$

Then s can be written in the form

$$s(x,\dot{x},t) = \frac{\partial\sigma(x,t)}{\partial t} + \sum_{l=1}^{n} \frac{\partial\sigma(x,t)}{\partial x_l}\dot{x}_l = \frac{d\sigma(x,t)}{dt},$$

which means that a necessary condition that $\wedge_k(f) = \wedge_k(h)$ is that there exists a twice continuously differentiable function $\sigma(x,t)$ of $n+1$ independent variables (x,t) such that

$$h = f + \frac{d\sigma(x,t)}{dt} \tag{1.1.7}$$

10 Conversely, if there exists a twice continuously differentiable function $\sigma(x,t)$ and if f and h are connected by the relation (1.1.7), then $\wedge_k(f) = \wedge_k(h)$ identically, $k = 1,\ldots,n$. This assertion could also have been proved starting from the original problem of the Calculus of Variations.

We proceed to derive the canonical equations of Hamilton. This is done by means of the 'Legendre transformation'. We set

$$y_k = f_{\dot{x}_k}(x,\dot{x},t),\ k = 1,\ldots,n, \tag{1.1.8}$$

and consider $f_{\dot{x}_k}$ as functions of $\dot{x} = (\dot{x}_1,\ldots,\dot{x}_n)$. If we suppose that the Jacobian $|f_{\dot{x}_k\dot{x}_l}| \neq 0$ everywhere in the region under consideration, then we can solve the system of equations (1.1.8) locally and determine $\dot{x}_k$ as functions of $2n+1$ independent variables (x,y,t). If f satisfies the Euler-Lagrange equations

$$f_{x_k} - \frac{d}{dt}f_{\dot{x}_k} = 0, \quad k = 1,\ldots,n,$$

then it follows that $\dot{y}_k = f_{x_k}$, $k = 1,\ldots,n$. Therefore, we obtain by the substitution (1.1.8) a system of $2n$ differential equations of the first order,

$$\dot{x}_k = \dot{x}_k(x,y,t),$$

$$\dot{y}_k = f_{x_k}(x, \dot{x}, t) = f_{x_k}(x, \dot{x}(x, y, t), t), \ k = 1, \ldots, n.$$

Thus the Euler-Lagrange system of n equations of the second order has been reduced to a system of $2n$ equations of the first order.

This system of differential equations can be written in a different 11
way as follows. Consider the function $E(x, y, \dot{x}, t)$ of $3n + 1$ independent real variables, defined by

$$E(x, y, \dot{x}, t) = \sum_{k=1}^{n} \dot{x}_k y_k - f(x, \dot{x}, t); \tag{1.1.9}$$

this is twice continuously differentiable in all the variables. Then

$$dE = \sum_{k=1}^{n} (\dot{x}_k dy_k + y_k d\dot{x}_k) - \sum_{k=1}^{n} (f_{x_k} dx_k + f_{\dot{x}_k} d\dot{x}_k) - f_t dt.$$

Now if we assume that $y_k = f_{\dot{x}_k}$, $k = 1, \ldots, n$, and $|f_{\dot{x}_k \dot{x}_1}| \neq 0$ everywhere, then

$$\sum_{k=1}^{n} (y_k d\dot{x}_k - f_{\dot{x}_k} d\dot{x}_k) = 0,$$

and hence

$$dE = \sum_{k=1}^{n} (\dot{x}_k dy_k - f_{x_k} dx_k) - f_t dt,$$

which means that E is a function $E(x, y, t)$, twice continuously differentiable in the $2n + 1$ independent variables (x, y, t). Therefore we get

$$\frac{\partial E}{\partial x_k} = -f_{x_k}, \quad \frac{\partial E}{\partial y_k} = \dot{x}_k \quad \text{and} \quad \frac{\partial E}{\partial t} = -f_t.$$

If now f satisfies the Euler-Lagrange equations, so that

$$f_{x_k} = \frac{d}{dt} f_{\dot{x}_k} = \dot{y}_k, \ k = 1, \ldots, n,$$

then we get a system of $2n$ differential equations satisfied by E:

$$\dot{x}_k = E_{y_k}, \dot{y}_k = -E_{x_k}, \ k = 1, \ldots, n. \tag{1.1.10}$$

where $E_{x_k} = \dfrac{\partial E(x,y,t)}{\partial x_k}$ and $E_{y_k} = \dfrac{\partial E(x,y,t)}{\partial y_k}$. These equations are 12
called the *canonical equations of Hamilton*. If we assume that $|y_{k\dot{x}_1}| = |f_{\dot{x}_k\dot{x}_1}| \neq 0$, then we see from the relation $y_k = f_{\dot{x}_k}$ (considering $f_{\dot{x}_k}$ as a function of $\dot{x}$) that

$$dy_k = \sum_{l=1}^{n} f_{\dot{x}_k\dot{x}_1} d\dot{x}_1.$$

Since $d\dot{x}_k = \sum\limits_{l=1}^{n} E_{y_k y_l} dy_1$, it follows that $(f_{\dot{x}_k\dot{x}_1})$ and $(E_{y_k y_l})$ are matrices which are inverses of each other. We have $E_{y_k y_l} = \dot{x}_{k y_l}$ and hence $|\dot{x}_{k y_l}| \neq 0$. Conversely, suppose that $E = E(x,y,t)$ is 0 given twice continuously differentiable function of all the $3n+1$ independent variables and $|E_{y_k y_l}| \neq 0$ everywhere, then the system of differential equations

$$\dot{x}_k = E_{y_k},\ \dot{y}_k = -E_{x_k},\ k = 1,\ldots,n,$$

can be reduced to the system of Euler-Lagrange equations. To see this we put

$$f(x,y,\dot{x},t) = \sum_{k=1}^{n} \dot{x}_k y_k - E(x,y,t) \tag{1.1.11}$$

and consider f as a function of $3n+1$ independent variables. Then

$$df = -\sum_{k=1}^{n}(E_{x_k}dx_k + E_{y_k}dy_k) - E_t dt + \sum_{k=1}^{n}(\dot{x}_k dy_k + y_k d\dot{x}_k).$$

13 If the equations (1.1.10) are satisfied, then

$$df = -\sum_{k=1}^{n} E_{x_k}dx_k - E_t dt + \sum_{k=1}^{n} y_k d\dot{x}_k,$$

and if $|E_{y_k y_1}| \neq 0$, then we can solve locally for y_l as a function of $\dot{x} = (\dot{x}_1,\ldots,\dot{x}_n)$. Substituting this in the expression (1.1.11) for f, we may consider f as a function of $2n+1$ independent variables $(x,\dot{x},t)$. Consequently,

$$f_{x_k} = -E_{x_k},\ f_{\dot{x}_k} = y_k,\ k = 1,\ldots,n,$$

which, together, with the system of equations $\dot{y}_k = -E_{x_k}$ of (1.1.10), implies that

$$f_{x_k} - \frac{d}{dt} f_{\dot{x}_k} = 0, \quad k = 1, \dots, n.$$

2 The transformation theory of Hamiltonian equations

We have shown in § 1 that the system of Hamiltonian differential equations can be obtained from the Euler-Lagrange differential equations (by means of the Legendre transformation) and conversely. We shall now show that the Hamiltonian equations can also be obtained directly from the variational problem, without using the Legendre transformation. This is done in the following way. Suppose that a twice continuously differentiable function $E(x, y, t)$ of $2n + 1$ independent real variables (x, y, t) is given. We generalize slightly and consider the function f of $4n + 1$ independent variables $(x, y, \dot{x}, \dot{y}, t)$ defined by **14**

$$f(x, y, \dot{x}, \dot{y}, t) = \sum_{r=1}^{n} \dot{x}_r y_r - E(x, y, t). \tag{1.2.1}$$

(The variable $\dot{y} = (\dot{y}_1, \dots, \dot{y}_n)$) does not really appear on the right hand side.) It is clear that the function f is twice continuously differentiable in all its variables. Then the Lagrangian derivatives of f, calculated with respect to $(x_k, \dot{x}_k)$ and $(y_k, \dot{y}_k)$ and denoted by $\wedge_{x_k}(f)$ and $\wedge_{y_k}(f)$ respectively, are given by

$$\begin{aligned} \wedge_{x_k}(f) &\equiv f_{x_k} - \frac{d}{dt} f_{\dot{x}_k} = -E_{x_k} - \dot{y}_k, \\ \wedge_{y_k}(f) &\equiv f_{y_k} - \frac{d}{dt} f_{\dot{y}_k} = \dot{x}_k - E_{y_k}, \ k = 1, \dots, n. \end{aligned} \tag{1.2.2}$$

Hence the Euler-Lagrange equations for f become

$$-E_{x_k} - \dot{y}_k = 0, \quad \dot{x}_k - E_{y_k} = 0, \quad k = 1, \dots, n, \tag{1.2.3}$$

which are precisely the Hamiltonian equations. Conversely the system (1.2.3) of equations implies that f satisfies the Euler-Lagrange equations. Thus the Hamiltonian equations can be considered as necessary conditions for the existence of a twice continuously differentiable function $(x(t), y(t)) = (x_1(t), \dots, x_n(t), y_1(t), \dots, y_n(t))$ which is a solution of the problem of minimising the integral

$$\int_{t_1}^{t_2} f(x, y, \dot{x}, \dot{y}, t)dt$$

with the prescribed initial conditions $(x(t_1), y(t_1))$ and $(x(t_2), y(t_2))$.

15 We proceed to discuss the transformation theory of the Hamiltonian equations. We consider $2n$ new independent real variables $\xi = (\xi_1, \dots, \xi_n)$ and $\eta = (\eta_1, \dots, \eta_n)$ and the transformation

$$x_k = \varphi_k(\xi, \eta, t),\ y_k = \psi_k(\xi, \eta, t),\ k = 1, \dots, n, \tag{1.2.4}$$

where φ_k and ψ_k are twice continuously differentiable functions of the $2n + 1$ independent variables (ξ, η, t) with the Jacobian $\dfrac{\partial(\varphi, \psi)}{\partial(x, y)}$ non-vanishing. In general such a transformation does not leave invariant the Euler-Lagrange equations in the Hamiltonian form:

$$E_{x_k} = -\dot{y}_k,\ E_{y_k} = \dot{x}_k,\ k = 1, \dots, n. \tag{1.2.5}$$

We wish to investigate conditions under which a transformation of the type (1.2.4) leaves the equations in the Hamiltonian form (1.2.5) invariant. For this purpose we consider the function f of $4n + 1$ independent variables $(x, y, \dot{x}, \dot{y}, t)$ defined by (1.2.1):

$$f(x, y, \dot{x}, \dot{y}, t) = \sum_{r=1}^{n} \dot{x}_r y_r - E(x, y, t).$$

We have seen that the Lagrangian derivatives of f, computed formally relative to $(x_k, \dot{x}_k)$ and $(y_k, \dot{y}_k)$ respectively, are

$$\wedge_{x_k(f)} = -E_{x_k} - \dot{y}_k \quad \text{and} \quad \wedge_{y_k}(f) = \dot{x}_k - E_{y_k}, k = 1, \dots, n.$$

These are in the Hamiltonian form. We substitute $\varphi_k(\xi, \eta, t)$ and $\psi_k(\xi, \eta, t)$ for x_k and y_k respectively in the expression (1.2.1) for f and we consider f as a function of the $4n+1$ new independent variables $(\xi, \eta, \dot{\xi}, \dot{\eta}, t)$. **16**

We then obtain the Lagrangian derivatives in the new variables:

$$\wedge_{\xi k}(f) \equiv f_{\xi_k} - \frac{d}{dt} f_{\dot{x}_k}, \wedge_{\eta_k}(f) \equiv f_{\eta,k} - \frac{d}{dt} f_{\dot{\eta}_k}, k = 1, \ldots, n.$$

Now if the transformation of the variables from (ξ, η) to (x, y) is to leave invariant the Lagrangian derivative in the Hamiltonian form, we have to determine a twice continuously differentiable function $\mathbb{E} = \mathbb{E}(\xi, \eta, t)$ of the $2n + 1$ independent real variables (ξ, η, t) such that

$$\wedge_{\xi_k}(f) = -\mathbb{E}_{\xi_k} - \dot{\eta}_k,\ \wedge_{\eta_k}(f) = \dot{\xi}_k - \mathbb{E}_{\eta_k},\ k = 1, \ldots, n. \tag{1.2.6}$$

On the other hand, if we consider the function $h = h(\xi, \eta, \dot{\xi}, \dot{\eta}, t)$ of the $4n + 1$ independent variables $(\xi, \eta, \dot{\xi}, \dot{\eta}, t)$ defined by

$$h(\xi, \eta, \dot{\xi}, \dot{\eta}, t) = \sum_{k=1}^{n} \dot{\xi}_n \eta_k - \mathbb{E}(\xi, \eta, t), \tag{1.2.7}$$

then the Lagrangian derivatives of h are again given by

$$\wedge_{\xi_k}(h) = -\mathbb{E}_{\xi_k} - \dot{\eta}_k, \wedge_{\eta_k}(h) = \dot{\xi}_k - \mathbb{E}_{\eta_k}, k = 1, \ldots, n. \tag{1.2.8}$$

The systems (1.2.6) and (1.2.8) together mean that the function $s = f - h$ is twice continuously differentiable in the variables $(\xi, \eta, \dot{\xi}, \dot{\eta}, t)$ and satisfies the Euler-Lagrange equations

$$\wedge_{\xi_k}(s) = 0,\ \wedge_{\eta_k}(s) = 0,\ k = 1, \ldots, n.$$

We have shown in § 1 that in such a case there exists a twice continuously differentiable function $\sigma = \sigma(\xi, \eta, t)$ of the $2n + 1$ independent **17**
variables (ξ, η, t) such that $s = \dfrac{d}{dt}\sigma(\xi, \eta, t)$. This means that $f = h + \dfrac{d\sigma}{dt}$ and hence, f considered as a function of the variables $(\xi, \eta, \dot{\xi}, \dot{\eta}, t)$ has the form

$$f(\xi, \eta, \dot{\xi}, \dot{\eta}, t) = \sum_{k=1}^{n} \dot{\xi}_k \eta_k - \mathbb{E}(\xi, \eta, t) + \frac{d}{dt}\sigma(\xi, \eta, t). \tag{1.2.9}$$

If we denote by σ_t, σ_{ξ_k}, σ_{η_k} the partial derivatives of σ with respect to t, ξ_k and η_k respectively, then we have

$$\frac{d}{dt}\sigma(\xi,\eta,t) = \sum_{k=1}^{n}(\sigma_{\xi_k}\dot{\xi}_k + \sigma_{\eta_k}\dot{\eta}_k) + \sigma_t.$$

Since $\dot{x}_r = \sum\limits_{k=1}^{n}(x_{r\xi_k}\dot{\xi}_k + x_{r\eta_k}\dot{\eta}_k) + x_{r_t}$, the expression (1.2.1) for f becomes

$$f = -E(x(\xi,\eta,t), y(\xi,\eta,t), t) + \sum_{r=1}^{n}((x_{r\xi_k}\dot{\xi}_k + x_{r\eta_k}\dot{\eta}_k) + x_{rt})y_r.$$

Then we get

$$\begin{aligned}\frac{d}{dt}\sigma(\xi,\eta,t) = f - h = \mathbb{E}(\xi,\eta,t) - E(x,y,t) - \sum_{k=1}^{n}\dot{\xi}_k\eta_k + \\ + \sum_{r=1}^{n}\left(\sum_{k=1}^{n}(x_{r\xi_k}\dot{\xi}_k + x_{r\eta_k}\dot{\eta}_k) + x_{rt}\right)y_r,\end{aligned}$$

and therefore, comparing the coefficients of $\dot{\xi}_k$ and $\dot{\eta}_k$ and the remaining terms, we have

$$\begin{aligned}\sigma_{\xi_k} &= \sum_{r=1}^{n} x_{r\xi_k}y_r - \eta_k,\\ \sigma_{\eta_k} &= \sum_{r=1}^{n} x_{r\eta_k}y_r,\ k = 1,\dots,n, \qquad (1.2.10)\\ \text{and}\qquad \sigma_t &= \sum_{r=1}^{n} x_{rt}y_r + \mathbb{E}(\xi,\eta,t) - E(x,y,t).\end{aligned}$$

18 The function $\mathbb{E}(\xi,\eta,t)$ is therefore determined by the last identity if the function $\sigma(\xi,\eta,t)$ is known. However, the first two identities in (1.2.10) give the partial derivatives of σ with respect to ξ_k and η_k. Hence a necessary and sufficient condition that there exist a twice continuously differentiable function σ of $2n + 1$ independent variables (ξ,η,t) with

$\dfrac{\partial\sigma}{\partial\xi_k}$ and $\dfrac{\partial\sigma}{\partial\eta_k}$ given by the first two equations in (1.2.10) is that σ satisfy the integrability conditions:

$$\sigma_{\xi_k\xi_1} = \sigma_{\xi_1\xi_k}, \sigma_{\xi_k\eta_1} = \sigma_{\eta_1\xi_k}, \sigma_{\eta_k\eta_1} = \sigma_{\eta_1\eta_k}, k, l = 1, \ldots, n.$$

Using the expressions for σ_{ξ_k} and σ_{η_1} from (1.2.10) the integrability conditions become

$$\begin{aligned}
\left(\sum_{r=1}^{n} x_{r\xi_k} y_r - \eta_k\right)_{\xi_1} &= \left(\sum_{r=1}^{n} x_{r\xi_l} y_r - \eta_1\right)_{\xi_k},\\
\left(\sum_{r=1}^{n} x_{r\xi_k} y_r - \eta_k\right)_{\eta_1} &= \left(\sum_{r=1}^{n} x_{r\eta_1} y_r\right)_{\xi_k},\\
\left(\sum_{r=1}^{n} x_{r\eta_k} y_r\right)_{\eta_1} &= \left(\sum_{r=1}^{n} x_{r\eta_1} y_r\right)_{\eta_k}.
\end{aligned}$$

Since x_r is twice continuously differentiable in all the variables (ξ, η, t), we have

$$x_{r\xi_k\xi_1} = x_{r\xi_1\xi_k}, x_{r\xi_k\eta_1} = x_{r\eta_1\xi_k}, x_{r\eta_k\eta_1} = x_{r\eta_1\eta_k},$$

$k, l = 1, \ldots, n,$ and we get

$$\begin{aligned}
\sum_{r=1}^{n} x_{r\xi_k} y_{r\xi_1} &= \sum_{r=1}^{n} x_{r\xi_1} y_{r\xi_k},\\
\sum_{r=1}^{n} x_{r\xi_k} y_{r\eta_1} - \delta_{kl} &= \sum_{r=1}^{n} x_{r\eta_1} y_{r\xi_k}, \qquad (1.2.11)\\
\sum_{r=1}^{n} x_{r\eta_k} y_{r\eta_1} &= \sum_{r=1}^{n} x_{r\eta_1} y_{r\eta_k},
\end{aligned}$$

$k, l = 1, \ldots, n,$ where $\delta_{kk} = 1$ and $\delta_{kl} = 0$, $k \neq 1$. These integrability **19**
conditions can best be written in matrix form. Denoting by A, B, C, D respectively the n-rowed square matrices

$$A = (x_{k\xi_1}), B = (x_{k\eta_1}), C = (y_{k\xi_1}), D = (y_{k\eta_1}),$$

we can write the equations (1.2.11) in the form

$$A'C = C'A,\ A'D - E = C'B,\ B'D = D'B, \tag{1.2.12}$$

where E is the n-rowed unit matrix (δ_{kl}). We denote by M the $2n$-rowed matrix

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

M is precisely the Jacobian matrix of the transformation from (ξ, η) to (x, y) given by $x_k = \varphi_k(\xi, \eta, t)$, $y_k = \psi_k(\xi, \eta, t)$, $k = 1, \ldots, n$. Denoting by J the $2n$-rowed matrix

$$J = \begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix},$$

20 where 0 stands for the n-rowed zero-matrix, the integrability conditions (1.2.12) can be condensed into the single condition:

$$M'JM = J. \tag{1.2.13}$$

A $2n$-rowed matrix M satisfying the condition (1.2.13) is called a *symplectic matrix*. We observe that $J' = -J$ and that $J^2 = -\begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}$, which shows that J itself is a symplectic matrix. We recall that the symplectic matrices form a group under matrix multiplication; this group is called the real symplectic group.

Thus the integrability conditions expressed by (1.2.13) state that the Jacobian matrix of the transformation from (ξ, η) to (x, y) defined by (1.2.4) is symplectic. Since the integrability conditions are derived from the first two equations in (1.2.10) which are independent of the function $E(x, y, t)$, it follows that the condition $M'JM = J$ is also independent of the choice of the function $E(x, y, t)$ in the expression (1.2.1) for f.

A transformation $x_k = \varphi_k(\xi, \eta, t)$, $y_k = \psi_k(\xi, \eta, t)$ where $\varphi_k, \psi_k, k = 1, \ldots, n$, are twice continuously differentiable functions of the $2n + 1$ independent variables (ξ, η, t) such that the Jacobian matrix M of the transformation is non-singular, is called *canonical* if the matrix M is symplectic.

Under a canonical transformation then the Hamiltonian equations preserve their form. We consider now the question of determining all canonical transformations, or, equivalently, the problem of solving the matrix equations $M'JM = J$ in M. We do this first under an additional 21
restriction. Suppose that $x_k = \varphi_k(\xi, \eta, t)$, $y_k = \psi_k(\xi, \eta, t)$, $k = 1, \ldots, n$, is a given transformation with the additional property that the determinant $B = |x_{k\eta_1}| \neq 0$; $|B|$ is precisely the Jacobian of the transformation $x_k = \varphi_k(\xi, \eta, t)$, $k = 1, \ldots, n$, where φ_k are considered as functions of the independent variables $\eta = (\eta_1, \ldots, \eta_n)$ alone. Now by the implicit function theorem we can slove locally for η_1 as functions of (x, ξ, t) : $\eta_1 = \eta_1(x, \xi, t)$ and substituting this in $y_k = \psi_k(\xi, \eta, t)$, we have $y_k = y_k(x, \xi, t)$. Since φ_k and ψ_k are twice continuously differentiable in all the variables (ξ, η, t), it follows that η_1, and hence y_1, are twice continuously differentiable functions of (x, ξ, t). Substituting $\eta_1 = \eta_1(x, \xi, t)$, $l = 1, \ldots, n$, in $\sigma = \sigma(\xi, \eta, t)$, we get a new function

$$W \equiv W(x, \xi, t) = \sigma(\xi, \eta(x, \xi, t), t),$$

which is again twice continuously differentiable in (x, ξ, t). Then we have the identity

$$\frac{d}{dt}\sigma(\xi, \eta, t) = \sum_{r=1}^{n} \dot{x}_r y_r - E(x, y, t) - \sum_{r=1}^{n} \dot{\xi}_r \eta_r + \mathbb{E}(\xi, \eta, t),$$

from which we obtain

$$\begin{aligned} \frac{d}{dt}W(x, \xi, t) = \sum_{r=1}^{n} \dot{x}_r y_r(x, \xi, t) - E(x, y(x, \xi, t), t) - \\ - \sum_{r=1}^{n} \dot{\xi}_r \eta_r(x, \xi, t) + \mathbb{E}(\xi, \eta(x, \xi, t), t). \end{aligned}$$

But since

$$\frac{d}{dt}W(x, \xi, t) = W_t + \sum_{k=1}^{n} (W_{x_k} \dot{x}_k + W_{\xi_k} \dot{\xi}_k),$$

we get, comparing the coefficients of $\dot{x}_k$, $\dot{\xi}_k$ and the remaining terms, 22

$$\begin{aligned} y_k = W_{x_k}, \eta_k = -W_{\xi_k},\ k = 1, \dots, n, \\ E(x, y, t) = \mathbb{E}(\xi, \eta, t) - W_t. \end{aligned} \tag{1.2.14}$$

Since $|B| = |x_{k\eta_1}| \neq 0$ and the matrix (η_{kx_1}) is the inverse of the matrix $(x_{k\eta_1})$, it follows that $|\eta_{kx_1}| \neq 0$. Therefore we have $|W_{\xi_k x_1}| = (-1)^n |\eta_{kx_1}| \neq 0$. Thus, given the transformation $x_k = \varphi_k(\xi, \eta, t)$, $y_k = \psi_k(\xi, \eta, t)$, $k = 1, \dots, n$, with $|x_{k\eta_1}| \neq 0$, there exists a twice continuously differentiable function $W = W(x, \xi, t)$ such that $|W_{\xi_k x_1}| \neq 0$. Conversely, suppose that we are given an arbitrary twice continuously differentiable function $W(x, \xi, t)$ of $2n + 1$ independent variables (x, ξ, t) with $|W_{\xi_k x_l}| \neq 0$. Then we set

$$\eta_k(x, \xi, t) = -W_{\xi_k}(x, \xi, t),\ y_k = W_{x_k}(x, \xi, t).$$

The first of these can be considered as a function of x only and since $|W_{\xi_k x_1}| = (-1)^n |\eta_{kx_1}| \neq 0$ by assumption, we can solve locally and obtain x_k as a function $\varphi_k(\xi, \eta, t)$ of (ξ, η, t). Substituting this in $y_k = W_{x_k}(x, \xi, t)$, we define the transformation

$$x_k = \varphi_k(\xi, \eta, t), y_k = \psi_k(\xi, \eta, t), k = 1, \dots, n.$$

Further, since the matrix $(x_{k\eta_1})$ is the inverse of the matrix $(\eta_{kx_1}) =$
23 $-(W_{\xi_k x_1})$, it follows that $|B| = |x_{k\eta_1}| \neq 0$.

Let us consider the identity transformation $x_k = \xi_k$, $y_k = \eta_k$, whose Jacobian matrix $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ is the $2n$-rowed identity matrix (evidently symplectic); here $B = 0$, so that the condition $|B| \neq 0$ is not satisfied and the theory above does not apply. However, even this case can be covered in the following way. We use the fact that nevertheless $|A| = |E| \neq 0$.

Suppose that we are given a transformation $x_k = \varphi_k(\xi, \eta, t), y_k = \psi_k(\xi, \eta, t)$ where φ_k, ψ_k are twice continuously differentiable functions of (ξ, η, t) with the Jacobian matrix $M = \dfrac{\partial(\varphi, \psi)}{\partial(\xi, \eta)}$ non-singular. Suppose that M has the additional property that $|A| = |x_{k\xi_1}| \neq 0$. We consider the transformation $\xi_k = \eta'_k$, $\eta_k = -\xi'_k$, $k = 1, \dots, n$, from the independent variables (ξ', η') to (ξ, η). The Jacobian matrix of this transformation is $\left(\begin{smallmatrix} 0 & E \\ -E & 0 \end{smallmatrix}\right) = J$ and J itself is a symplectic matrix. The Jacobian matrix of the composite transformation of the variables (ξ', η') to (x, y) defined by

$$x_k = \varphi_k(\xi, \eta, t) = \varphi_k(\eta', -\xi', t)$$

$$y_k = \psi_k(\xi, \eta, t) = \psi_k(\eta', -\xi', t), k = 1, \dots, n,$$

is the product matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix} = \begin{pmatrix} -B & A \\ -D & C \end{pmatrix}.$$

Since J is symplectic and the symplectic matrices form a group under matrix multiplication, we see that if $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ is symplectic, then **24**
$\left(\begin{smallmatrix} -B & A \\ -D & C \end{smallmatrix}\right)$ is also symplectic and conversely. Thus the Jacobian matrix of the composite of the two transformations is of the form previously considered since $|A| \neq 0$, and hence our argument can be applied to prove that there exists a twice continuously differentiable function $W' = W'(x, \xi', t)$ such that $|W'_{x_k \xi'_1}| \neq 0$ and $y_k = W'_{x_k}$, $\eta'_k = -W'_{\xi'_k}$, $k = 1, \dots, n$.

We remark that the transformation $x_k = \eta_k$, $y_k = -\xi_k$ whose Jacobian matrix is $\left(\begin{smallmatrix} 0 & E \\ -E & 0 \end{smallmatrix}\right) = J$ belongs to the type we have considered and moreover J is symplectic. This proves the existence of non-trivial canonical transformations with the property that $|B| = |x_{k\eta_1}| \neq 0$.

If both $|A|$ and $|B|$ are zero we can still proceed using the fact that a symplectic matrix can be expressed as a product of two symplectic matrices in each of which either the condition $|A| \neq 0$ or $|B| \neq 0$ is satisfied.

We now come to the partial differential equation of Hamilton-Jacobi. Consider the system of Hamiltonian differential equations

$$\dot{x}_k = E_{y_k}, \dot{y}_k = -E_{x_k},\ k = 1, \dots, n,$$

where $E = E(x, y, t)$ is a twice continuously differentiable function of the $2n + 1$ variables (x, y, t). Taking $f = \sum_{k=1}^{n} \dot{x}_k y_k - E$ we can write this in the form

$$\wedge_{x_k}(f) \equiv -\dot{y}_k - E_{x_k} = 0,\ \wedge_{y_k}(f) \equiv \dot{x}_k - E_{y_k} = 0,\ k = 1, \dots, n.$$

Let $x_k = \varphi_k(\xi, \eta, t)$, $y_k = \psi_k(\xi, \eta, t)$ be a transformation of the variables **25**
(ξ, η) to (x, y), where φ_k, ψ_k are twice continuously differentiable functions of all the variables (ξ, η, t) with Jacobian $\dfrac{\partial(\varphi, \psi)}{\partial(\xi, \eta)} \neq 0$. If this is a

canonical transformation, then there exists a twice continuously differentiable function $\mathbb{E} = \mathbb{E}(\xi, \eta, t)$ such that

$$\wedge_{\xi_k}(f) = -\dot{\eta}_k - \mathbb{E}_{\xi_k}, \wedge_{\eta_k}(f) = \dot{\xi}_k - \mathbb{E}_{\eta_k}, k = 1, \dots, n,$$

where f is considered as a function of the variables $(\xi, \eta, \dot{\xi}, \dot{\eta}, t)$. Since the Lagrangian derivatives are invariant under a canonical transformation, this would imply that $\wedge_{\xi_k}(f) = 0$, $\wedge_{\eta_k}(f) = 0$, $k = 1, \dots, n$.

Now suppose that the canonical transformation above: $x_k = \varphi_k(\xi, \eta, t)$, $y_k = \psi_k(\xi, \eta, t)$, is such that $\mathbb{E}(\xi, \eta, t) \equiv 0$; then the Hamiltonian differential equations take the trivial form $\dot{\xi}_k = 0$, $\dot{\eta}_k = 0$, which has a trivial solution $\xi_k =$ constant, $\eta_k =$ constant, $k = 1, \dots n$.

We have constructed a canonical transformation starting from a 'generating function', i.e. a twice continuously differentiable function $W = W(x, \eta, t)$ of the $2n + 1$ variables (x, ξ, t) with $|W_{x_k \xi_1}| \neq 0$. Then we have seen that W satisfies the relation $\mathbb{E}(\xi, \eta, t) = E(x, y, t) + W_t$. Hence, in order that $\mathbb{E}(\xi, \eta, t) \equiv 0$ it is necessary and sufficient that
26 $E(x, y, t) + W_t = 0$. We obtained the canonical transformation by defining $y_k = W_{x_k}$. Then this condition becomes

$$E(x, W_x, t) + W_t = 0, \tag{1.2.15}$$

and this is the *Hamilton-Jacobi partial differential equation* satisfied by W. Thus, if the canonical transformation constructed from a generating function W transforms the Hamiltonian differential equations into the trivial form $\dot{\xi}_k = 0$, $\dot{\eta}_k = 0$, $k = 1, \dots, n$, then W satisfies the Hamilton-Jacobi partial differential equation.

Conversely, suppose that W satisfies the Hamilton-Jacobi partial differential equation; then we obtain a canonical transformation in the following way. Define $\eta_k = -W_{\xi_k}, y_k = W_{x_k}$. Since $|W_{\xi_k x_l}| = |W_{x_k \xi_1}| \neq 0$, we can solve the equation $\eta_k = -W_{\xi_k}$ locally and express x_k as a function $\varphi_k(\xi, \eta, t)$ which, on substitution in $y_k = W_{x_k}$, gives $y_k = \psi_k(\xi, \eta, t)$. Moreover, since under this transformation $\mathbb{E}(\xi, \eta, t) = E(x, y, t) + W_t = E(x, W_x, t) + W_t = 0$, it follows that the transformation thus obtained reduces the Hamiltonian system of differential equations into the trivial form $\dot{\xi}_k = 0$, $\dot{\eta}_k = 0$.

3 Cauchy's theorem on the existence of solutions of a system of ordinary differential equations

Suppose that we are given m real-valued functions $f_k = f_k(x_1, \ldots, x_m)$, $k = 1, \ldots, m$, of m independent real variables $(x_1, \ldots, x_m)$ and m real numbers $\xi_1, \ldots, \xi_m$. Let now $x_k = x_k(t)$ be m real-valued functions of a real variable t in some interval; denote by $\dot{x}_k$ the derivative $\frac{d}{dt}x_k(t)$. We 27
shall consider the system of ordinary differential equations

$$\dot{x}_k = f_k(x_1, \ldots, x_m),\ k = 1, \ldots, m,$$

in the m unknown functions $x_k(t)$ taking the initial values ξ_k at the point $t = \tau : x_k(\tau) = \xi_k$. It is well known that if f_k are, for instance, Hölder continuous in a real neighbourhood of the point $(\xi_1, \ldots, \xi_m)$ in m-dimensional Euclidean space, then there exists a solution x_k of the system of differential equations in a real neighbourhood of τ, satisfying the initial condition $x_k(\tau) = \xi_k$. We shall consider the system of differential equations in the complex domain and seek complex solutions x_k. More precisely, let $\xi_1, \ldots, \xi_m$ be m given complex numbers. We shall assume that the f_k are complex valued regular analytic functions of m independent complex variables $(x_1, \ldots, x_m)$ in a complex neighbourhood of $(\xi_1, \ldots, \xi_m)$:

$|x_k - \xi_k| < r_k$, $r_k > 0$, $k = 1, \ldots, m$. To simplify the notation we shall assume that the f_k are regular analytic functions in the region $|x_k - \xi_k| < r = \min(r_1, \ldots, r_m)$. Let us suppose further that there is a positive constant C such that $|f(x_1, \ldots, x_m)| \leq C$ for x in the region $|x_k - \xi_k| < r$. We shall prove the following existence theorem due to Cauchy.

Theorem. *If f_k are regular analytic functions of m complex variables $(x_1, \ldots, x_m)$ in a complex neighbourhood $|x_k - \xi_k| < r$ of the point $(\xi_1, \ldots, \xi_m)$ and $|f_k| \leq C$ in this region, then the system of differential equations*

$$\dot{x}_k = f_k(x_1, \ldots, x_m),\ k = 1, \ldots, m, \tag{1.3.1}$$

has a solution $x_k = x_k(t)$ in the complex neighbourhood 28

$$|t - \tau| < r/(m+1)C \tag{1.3.2}$$

of the point τ, such that the $x_k(t)$ are regular analytic functions of the variable t in this region with $x_k(\tau) = \xi_k$ and

$$|x_k(t) - \xi_k| < r, \ k = 1, \ldots, m$$

in the region (1.3.2).

Proof. The idea used by Cauchy is the following. One writes the x_k as power-series with undertermined coefficients, inserts these into the differential equations (1.3.1) and equates the coefficients on both sides; the coefficients of the power-series for x_k are now determined and one then proves the convergence of the resulting power-series by the method of majorization. We shall first of all simplify the notation in the following way. Define new variables $x_1^*, \ldots, x_m^*$ and t^* by means of the substitutions

$$x_k^* = (x_k - \xi_k)/r, \ k = 1, \ldots, m; t^* = \frac{c}{r}(t - \tau),$$

$$\text{i.e.} \qquad x_k = r x_k^* + \xi_k, \quad \text{and} \quad t = \frac{r}{C} t^* + \tau.$$

Then $|x_k^*| < 1$ and $|t^*| < 1/(m+1)$ for $(x_1, \ldots, x_m, t)$ in the region $|x_k - \xi_k| < r$ and $|t - \tau| < r/(m+1)C$. Now the system of differential equations (1.3.1) becomes

$$c\frac{dx_k^*}{dt^*} = f_k(r x_1^* + \xi_1 \ldots, r x_m^* + \xi_m).$$

29 Setting $f_k^*(x_1^*, \ldots, x_m^*) = \dfrac{1}{C} f_k(r x_1^* + \xi_1, \ldots, r x_m^* + \xi_m)$, this takes the form

$$\frac{dx_k^*}{dt^*} = f_k^*(x_1^*, \ldots, x_m^*), \ k = 1, \ldots, m, \tag{1.3.3}$$

where f_k^* are regular analytic functions of the new complex variables $(x_1^*, \ldots, x_m^*)$ in the region $|x_k^*| < 1$ and further, $|f_k^*| \leq 1$ in this region. This is again of the form (1.3.1). Now the statement of the theorem reads as follows: there exists a complex regular analytic solution $x_k^* = x_k^*(t^*)$ in the complex region $|t^*| < 1/(m+1)$ of the system (1.3.3) with the initial condition $x_k^*(0) = 0$ and with $|x_k^*(t^*)| < 1$ for $|t^*| < 1/(m+1)$. It is clear that every solution $x_k^*(t^*)$ of this problem gives a solution of the original problem (1.3.1) and vice versa. Hence it is enough to consider the system (1.3.1) in the situation in which $\xi_k = 0$, $\tau = 0$, f_k regular analytic in the complex region $|x_k| < 1$, and $|f_k| \leq 1$.

We shall, first of all, construct a formal solution of the system (1.3.1) with the initial condition $x_k(0) = 0$. Since we seek regular analytic solutions $x_k = x_k(t)$ with $x_k(0) = 0$, we shall consider the formal power-series

$$x_k = x_k(t) = \sum_{n=1}^{\infty} \alpha_{k,n} t^n, \tag{1.3.4}$$

where the coefficients $\alpha_{k,n}$ are complex numbers, to be determined. We introduce the following notation in order to simplify the writing. If $\varphi = \sum\limits_{l=0}^{\infty} c_l t^l$ is a formal power-series in one variable with complex co- **30**
efficients, for each integer $n \geq 0$ we shall denote the partial sum $\sum\limits_{l=0}^{n} c_l t^l$ by φ_n, and the coefficient of t^n by $(\varphi)_n$. It is clear that $(\varphi_n)_n = (\varphi)_n$. Further, if ψ is another formal power series, then we have $(\varphi \pm \psi)_n = \varphi_n \pm \psi_n$ and $(\varphi\psi)_n = (\varphi_n \psi_n)_n$. Since each f_k is a complex regular analytic function of the variables $(x_1, \ldots, x_m)$, is has a power-series expansion with complex coefficients:

$$f_k = \sum_{l_1,\ldots,l_m=0}^{\infty} a_{k,l_1\ldots l_m} x_1^{l_1} \ldots x_m^{l_m}. \tag{1.3.5}$$

For the moment we shall not be interested in the convergnece of this series. Substituting (1.3.4) for $x_k(t)$ and $\sum\limits_{n=0}^{\infty} (n+1)\alpha_{k,n+1} t^n$ for $\dot{x}_k(t)$ in the differential equations $\dot{x}_k = f_k(x_1, \ldots, x_m)$ and comparing the coefficients of t^n on the two sides, we obtain, using (1.3.5),

$$(n+1)\alpha_{k,n+1} = \sum_{l_1\ldots l_m=0}^{\infty} a_{k,l_1\ldots l_m} (x_1^{l_1} \ldots x_m^{l_m})_n. \tag{1.3.6}$$

We observe that the power-series for $x_k(t)$ contains no constant term and consequently, there is no contribution to the term t^n on the right side in

(1.3.6) if $l_1 + \ldots + l_m > n$; hence, for each n, the right side of (1.3.6) contains only finitely many terms. Then, with the notation introduced earlier, (1.3.6) becomes

$$(n+1)\alpha_{k,n+1} = \sum_{l_1 \ldots l_m = 0}^{\infty} a_{k,l_1 \ldots l_m}((x_{1n})^{l_1} \ldots (x_{mn})^{l_m})_n \tag{1.3.7}$$

31 This is a recurrence formula for determining the coefficients $\alpha_{k,n}$, $k = 1, \ldots, m$; $n = 1, 2, \ldots$. We show now by induction on n that the $\alpha_{k,n}$ are polynomials in $a_{r,l_1 \ldots l_m}$ with non-negative rational coefficients. In fact, we have $\alpha_{k,1} = a_{k,0 \ldots 0}$ for $k = 1, \ldots, m$, so that we can start the induction. Suppose that $\alpha_{k,1}, \ldots, \alpha_{k,n}$, $k = 1, \ldots, m$, have already been determined as polynomials in $a_{r,l_1 \ldots l_m}$ with non-negative rational coefficients. Since each $x_{k,n} = \sum_{q=1}^{n} \alpha_{k,q} t^q$, and $l_1, \ldots, l_m$ are non-negative integers, it follows that $((x_{1n})^{l_1} \ldots (x_{mn})^{l_m})_n$ is a polynomial in $a_{r,p_1,\ldots,p_m}$ and hence, by (1.3.7), so is $\alpha_{k,n+1}$. Thus the coefficients $\alpha_{k,n}$ in the formal power-series expansion (1.3.4) for x_k are determined.

Next we shall prove the convergence of the formal power-series (1.3.4) for $x_k = x_k(t)$. For this we make use of the method of majorants and this idea is due to Cauchy. Suppose that

$$f = \sum_{l_1 \ldots l_m = 0}^{\infty} a_{l_1 \ldots l_m} x_1^{l_1} \ldots x_m^{l_m}, \; g = \sum_{l_1 \ldots l_m = 0}^{\infty} b_{l_1 \ldots l_m} x_1^{l_1} \ldots x_m^{l_m}$$

are two formal power-series in m variables with the coefficients $a_{l_1 \ldots l_m}$ complex and $b_{l_1 \ldots l_m}$ non-negative real numbers. We shall say that f *is majorized by* g (or g is a *majorant* of f) if

$$|a_{l_1 \ldots l_m}| \leq b_{l_1 \ldots l_m}, \; l_1, \ldots, l_m = 0, \; 1, \ldots$$

and we denote this by $f \prec g$ or $g \succ f$. If f_k and g_k are formal power-
32 series with $f_k \prec g_k$, then the system of differential equations

$$\dot{y}_k = g_k(y_1, \ldots, y_m), \; k = 1, \ldots, m,$$

where $g_k = \sum_{l_1 \ldots l_m = 0}^{\infty} b_{k,l_1 \ldots l_m} y_1^{l_1} \ldots y_m^{l_m}$, is called a *majorant system*. This system of differential equations with the initial condition $y_k(0) = 0$,

$k = 1, \ldots, m$, can be solved as above and one obtains a formal solution $y_k(t)$ in a formal power-series in one variable t:

$$y_k(t) = \sum_{n=1}^{\infty} \beta_{k,n} t^n, \ k = 1, \ldots, m.$$

Once again the coefficients $\beta_{k,n}$ are determined by means of the recurrence formula

$$(n+1)\beta_{k,n+1} = \sum_{l_1 \ldots l_m = 0}^{\infty} b_{k,l_1 \ldots l_m}((y_{1n})^{l_1} \ldots (y_{mn})^{l_m})_n. \tag{1.3.8}$$

As before one can show by induction that the $\beta_{k,n}$ are polynomials in $b_{r,l_1 \ldots l_m}$ with non-negative rational coefficients. Since $b_{r,l_1 \ldots l_m}$ are themselves non-negative, we see that the $\beta_{k,n}$ are non-negative.

Now we shall show that if $y_k = y_k(t)$ is a solution of a majorant system

$$\dot{y}_k = g_k(y_1, \ldots, y_m), \ k = 1, \ldots, m, \tag{1.3.9}$$

of the system (1.3.1), with initial conditions $y_k(0) = 0$, then $x_k \prec y_k$, $k = 1, \ldots, m$, as power-series in the variable t. In other words, we show that

$$|\alpha_{k,n}| \leq \beta_{k,n}, \ k = 1, \ldots, m; \ n = 1, 2, \ldots \tag{1.3.10}$$

This is done by induction on n. Since for $n = 1$ we have **33**

$$|\alpha_{k,1}| = |a_{k,0\ldots0}| \leq b_{k,0\ldots0} = \beta_{k,1},$$

we can start the induction. Suppose that (1.3.10) has been proved for $n = 1, \ldots, q$; then by the recurrence formulas (1.3.7) and (1.3.8) we have

$$\begin{aligned}(q+1)|\alpha_{k,q+1}| &= \Big|\sum_{l_1 \ldots l_m = 0}^{\infty} a_{k,l_1 \ldots l_m}((x_{1q})^{l_1} \ldots (x_{mq})^{l_m})_q\Big| \\ &\leq \sum_{l_1 \ldots l_m = 0}^{\infty} |a_{k,l_1 \ldots l_m}| \, |((x_{1q})^{l_1} \ldots (x_{mq})^{l_m})_q|\end{aligned}$$

$$
\begin{aligned}
&\le \sum_{l_1 \ldots l_m = 0}^{\infty} b_{k, l_1 \ldots l_m} ((y_{1q})^{l_1} \ldots (y_{mq})^{l_m})_q \\
&= (q+1)\beta_{k,q+1}.
\end{aligned}
$$

This proves the assertion that $x_k \prec y_k$, $k = 1, \ldots m$.

Thus, in order to prove that the $x_k(t)$ are regular analytic functions of the variable t, it is enough to determine a suitable majorant system of differential equations (1.3.9) and solve it for $y_k(t)$, with initial condition $y_k(0) = 0$, as a power-series convergent in some region. This is done in the following way. If f_k are regular analytic functions of the complex variables $x_1, \ldots, x_m$ in a complex region $|x_k| < r_k$, $k = 1, \ldots, m$, then by Cauchy's integral formula we have

$$
a_{k, l_1 \ldots l_m} = \frac{1}{(2\pi i)^m} \int_{C_1} \cdots \int_{C_m} \frac{f_k(x_1, \ldots, x_m)}{x_1^{l_1+1} \ldots x_m^{l_m+1}} dx_1 \ldots dx_m,
$$

34 where C_k denotes the circle $|x_k| = \rho_k < r_k$, $k = 1, \ldots, m$. If $|f_k(x_1, \ldots, x_m)| \le M$, then it follows that

$$
|a_{k, l_1 \ldots l_m}| \le M / \rho_1^{l_1} \ldots \rho_m^{l_m}, l_1, \ldots, l_m = 0, 1, 2, \ldots
$$

Since in our case $M = 1$ and $r = 1$, we have $|a_{k, l_1 \ldots l_m}| \le 1$ for all $l_1, \ldots, l_m = 0, 1, 2, \ldots$. Hence we choose $b_{k, l_1 \ldots l_m} = 1$ for $k = 1, \ldots, m$ and all $l_1, \ldots, l_m = 0, 1, 2 \ldots$. Thus for each g_k, $k = 1, \ldots, m$, we take the power-series

$$
\sum_{l_1, \ldots, l_m = 0}^{\infty} y_1^{l_1} \ldots y_m^{l_m},
$$

which is the product of m geometric series with sum

$$
\frac{1}{(1-y_1) \cdots (1-y_m)}.
$$

This is independent of k and hence a majorant system for (1.3.1) is given by

$$
\dot{y}_k = \frac{1}{(1-y_1) \ldots (1-y_m)}, \ k = 1, \ldots, m,
$$

with the initial condition $y_k(0) = 0$. A solution of this system is given by the solution $y_1 = \ldots = y_m = y = y(t)$ of the single differential equation

$$\dot{y}(t) = 1/(1-y)^m,$$

with the initial condition $y(0) = 0$; integrating this we get **35**

$$y(t) = 1 - (1-(m+1)t)^{\frac{1}{m+1}}.$$

Expanding the right side as a binomial series we get

$$y(t) = \sum_{n=1}^{\infty} \binom{\frac{1}{m+1}}{n} (-1)^{n-1}(m+1)^n t^n.$$

The coefficients in this expansion are all positive and the power-series converges for $|t| < 1/(m+1)$. Since $y_k > x_k$, this proves that the power-seris (1.3.4) for $x_k(t)$ converges for $|t| < 1/(m+1)$, so that the $x_k(t)$ are regular analytic functions in this region of the complex t-plane.

It now remains only to show that $|x_k(t)| < 1$ in the region $|t| < 1/(m+1)$. This is an immediate consequence of the fact that for $|t| < 1/(m+1)$, we have

$$|x_k(t)| \leq y(|t|) = 1 - (1-(m+1)|t|)^{\frac{1}{m+1}} < 1.$$

This completes the proof of the theorem.

We have so far assumed that the f_k do not contain the variable t explicitly. However, the case in which the f_k are regular analytic functions of the $m+1$ complex variables $(x_1, \ldots, x_m, t)$ in the neighbourhood $|x_k - \xi_k| < r, |t-\tau| < r$ of the point $(\xi_1, \ldots, \xi_m, t)$ can be covered as follows. We take $x_{m+1} = t$ and consider the system of $m+1$ differential equations

$$\begin{aligned} \dot{x}_k &= f_k(x_1, \ldots, x_m, x_{m+1}),\ k = 1, \ldots, m, \\ \dot{x}_{m+1} &= 1, \end{aligned}$$

with initial conditions $x_k(\tau) = \xi_k, k = 1, \ldots, m$, and $x_{m+1}(\tau) = \tau$. Thus **36**
we obtain Cauchy's theorem in this more general case also.

For our applications to Celestial Mechanics we need consider only a real variable t. Let us suppose that $\tau, \xi_1, \dots, \xi_m$ are $m+1$ real numbers and that the f_k are real:

$$f_k = \sum_{l_1 \dots l_m = 0}^{\infty} a_{k,l_1 \dots l_m} x_1^{l_1} \dots x_m^{l_m},$$

with the coefficients $a_{k,l_1 \dots l_m}$ real; then by the recurrence formula (1.3.7), the coefficients $\alpha_{k,n}$ in the power-series expansion $x_k(t) = \sum_{n=0}^{\infty} \alpha_{k,n}(t-\tau)^n$ of the solution of the system of equations (1.3.1) with initial condition $x_k(\tau) = \xi_k$, being polynomials in $a_{r,l_1 \dots l_m}$ with non-negative coefficients, are themselves real.

Consider the half-open interval $t_1 \le t < t_2$ and suppose that $x_k = x_k(t)$ are real-valued regular analytic functions of the variable t in this interval. We have therefore a regular analytic curve in m-dimensional Euclidean space. Assume that $f_k = f_k(x_1, \dots, x_m)$ are regular analytic functions in a bounded closed point set D of m-dimensional Euclidean space containing this curve $x(t)$, $t_1 \le t < t_2$ and suppose that the functions $x_k = x_k(t)$ satisfy the differential equations $\dot{x}_k = f_k(x_1, \dots, x_m)$ in
37 the interval $t_1 \le t < t_2$. Then we claim that the solutions $x_k(t)$ of the system (1.3.1) which are regular analytic in the interval $t_1 \le t < t_2$ can be continued analytically as anlytic functions regular also at $t = t_2$. This can be proved in the following way. □

For every point $(\xi_1, \dots, \xi_m)$ of D there exists a complex neighbourhood $|x_k - \xi_k| < r_k$, $k = 1, \dots, m$, in which f_k is a regular analytic function of the m complex variables $(x_1, \dots, x_m)$. As $(\xi_1, \dots, \xi_m)$ runs through the point set D such neighbourhoods cover D. The union of all these neighbourhoods is an open point set F. Then by the idea of the proof of the Heine-Borel theorem we can choose a sufficiently small positive number r such that a finite union G of the neighbourhoods $|x_k - \xi_k| \le r$, $k = 1, \dots, m$ contains D and is contained in F. Since G is closed and bounded and the m functions f_k are regular analytic everywhere on G, it follows that each f_k is bounded on G. Therefore we can assume that $|f_k| \le C$ in the region $|x_k - \xi_k| \le r$, $k = 1, \dots, m$, where C is a positive constant independent of the point ξ in D.

Now take any real number τ in the interval $t_1 \leq t < t_2$. Then by Cauchy's existence theorem there exists a regular analytic solution $x^*(t) = (x_1^*(t), \ldots, x_m^*(t))$ in the complex region $|t-\tau| < r/(m+1)C$, taking the initial values $x_k^*(\tau) = \xi_k$. If we now take τ in $t_1 \leq t < t_2$ such that $|t_2 - \tau| < r/(m+1)C$ and $\xi_k = x_k(\tau)$, then it follows that the solution $x_k^*(t)$ is regular analytic at $t = t_2$, $k = 1, \ldots, m$. Since $x_k(t)$ and $x_k^*(t)$ are regular analytic functions of the variable t in the (connected) region 38
$|t-\tau| < r/(m+1)C$, $\tau \leq t < t_2$, and $x_k(\tau) = \xi_k = x_k^*(\tau)$, we conclude that $x_k(t) = x_k^*(t)$ for all t in this region. This shows that $x_k(t)$ can be continued analytically on the real interval $\tau \leq t \leq t_2$ so as to be regular at $t = t_2$.

In the following we shall be interested in applying the Cauchy existence theorem to a Hamiltonian system of differential equations. We observe that in this case the functions f_k in the system $\dot{x}_k = f_k(x_1, \ldots, x_m)$, $k = 1, \ldots, m$, are obtained starting from a single function $E = E(x, y, t)$ which is twice ontinuously differentiable in all its variables (x, y, t). (We have used the obvious notation $m = 2n$ and $x_1, \ldots, x_m$ stand for the $2n$ independent variables x, y). In fact, the functions f_k are the derivatives E_{y_k} and - E_{x_k}, and so in order to apply the Cauchy existence theorem we need estimates for - E_{x_k} and E_{y_k}. If the function E is a regular analytic function of its variables (x, y, t) in some complex region and is bounded by a constant M there, then one can obtain a bounded C for E_{x_k} and E_{y_k} in terms of M by using the Cauchy integral formula. Since the domain of existence of the solution depends on this constant C, it follows that this domain can be determined in terms of M itself in our case.

In order to make this more precise we begin with the following remark. Let $f(x)$ be a regular analytic function of are complex variable in the disc $|x| < r$ and let $|f(x)| \leq C$ in $|x| < r$. If ξ is any point in the disc, then by the Cauchy integral formula we have

$$f'(\xi) = \frac{1}{2\pi i} \int_\Gamma \frac{f(x)}{(x-\xi)^2} dx,$$

where Γ is a simple closed curve around ξ and contained in the disc 39
$|x| < r$. We however assume now that f is regular analytic in a larger disc $|x| < 2r$ and restrict ourselves to points ξ in the closed disc $|\xi| \leq r$.

We take for the curve Γ the circle $|x - \xi| = \rho$ where $0 < \rho < r$. Then we get, from the formula above, the estimate $|f'(\xi)| \leq C/r$.

Now we take $m = 2n$ and the $2n$ variables $(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_n)$. Let ξ_k, η_k, $k = 1, \ldots, n$, be $2n$ given complex constants. Suppose that $E = E(x, y)$ is a regular analytic function of the $2n$ complex variables (x, y) in the region $|x_k - \xi_k| < 2r$, $|y_k - \eta_k| < 2r$, $k = 1, \ldots, n$, and is independent of the variable t. Then, by the remarks made above, considering E as a function of the variables x_k and y_k in turn, it follows immediately that for all points (x, y) in the region $|x_k - \xi_k| \leq r, |y_k - \eta_k| \leq r, k = 1, \ldots, n$, we have the estimates $|Ey_k(x, y)| \leq C/r$, $| - E_{x_k}(x, y)| \leq C/r$. Consequently, by the existence theorem we see that if τ is a given complex number, then there exists a regular analytic solution $x_k = x_k(t)$, $y_k = y_k(t)$ of the Hamiltonian system of equations $\dot{x}_k = E_{y_k}$, $\dot{y}_k = -E_{x_k}$, $k = 1, \ldots, n$, in the complex neighbourhood $|t - \tau| < r^2/(m + 1)C$ and that the initial conditions $x_k(\tau) = \xi_k, y_k(\tau) = \eta_k$ are satisfied and $|x_k(t) - \xi_k| < r$, $|y_k(t) - \eta_k| < r$, $k = 1, \ldots, n$.

40 We remark that the case in which E is a regular analytic function of all the $2n + 1$ variables (x, y, t) can be considered exactly in the same way by taking t to be the $(2n + 1)^{\text{th}}$ variable z and the system of $2n + 1$ differential equations

$$\dot{x}_k = E_{y_k}, \dot{y}_k = -E_{x_k}, k = 1, \ldots, n, \dot{z} = 1,$$

with the initial conditions $x_k(\tau) = \xi_k$, $y_k(\tau) = \eta_k$ and $z(\tau) = \tau$. The remark on the continuation of the solution to the right-hand end-point of a real t-interval we made earlier is valid in this case also.

Chapter 2

The three-body problem : simple collisions

1 The n-body problem

We shall intordue the problem of n bodies in three-dimensional Euclidean space and study its singularities in the case $n = 3$. 41

Let n be an integer ≥ 2. (The case $n = 1$ will be seen to be trivial). Suppose that $P_1, \ldots, P_n$ are n point-masses in three-dimensional Euclidean space, with the rectangular cartesian coordinates of P_k denoted by (x_k, y_k, z_k), $k = 1, \ldots, n$. For simplicity we write q_k for any one of the three coordinates x_k, y_k, z_k, $k = 1, \ldots, n$, and q for any one of the $3n$ coordinates q_k. The distance r_{kl} between the points P_k and P_l is given by

$$r_{kl}^2 = (x_k - x_1)^2 + (y_k - y_1)^2 + (z_k - z_1)^2. \tag{2.1.1}$$

We shall suppose that P_k has a mass $m_k > 0$, $k = 1, \ldots, n$, and that $r_{kl} > 0$, $k \neq l$. Suppose that the n point-masses attract each other according to Newton's law of gravitation. Then we can write down the equations of motion of the system of n point-masses. For this we set

$$U = \sum_{1 \leq k < l \leq n} \frac{m_k m_1}{r_{kl}} \tag{2.1.2}$$

with defines the Newtonian gravitational potential of the system of n
42 point-masses P_k. The sum on the right in (2.1.2) contains $n(n-1)/2$ terms. We have assumed that the gravitation constant is 1 and this can always be done by choosing the unit of mass properly. Then the differential equations of motion of the system of n point-masses have the form

$$m_k\ddot{q}_k = U_{\ddot{q}_k},\ k = 1, \dots, n, \tag{2.1.3}$$

where q_k are considered as functions of the real (time) variable t and $\ddot{q}_k$ denotes the second derivative $\frac{d^2}{dt^2}q_k(t)$, while U_{q_k} denotes the partial derivative $\frac{\partial}{\partial q_k}U(x_1, \dots, z_n)$. This is a system of $3n$ ordinary differential equations of the second order in the $3n$ unknown functions $q = q(t)$ of the variable t; we can write them symbolically in the form

$$m\ddot{q} = U_q, \tag{2.1.4}$$

where m denotes the mass associated with q. We can also write this as a system of $6n$ ordinary differential equations of the first order by introducing the velocity components $v_k = \dot{q}_k = \frac{d}{dt}q_k(t)$:

$$\dot{q} = v,\ m\dot{v} = U_q. \tag{2.1.5}$$

These are $6n$ ordinary differential equations in $6n$ unknown functions $q_k(t)$, $v_k(t)$ of the variable t. We shall start from the initial time $t = \tau$, a real number, and we prescribe the initial values $q_k(\tau) \equiv q_{k\tau}$ for $q(t)$ at
43 $t = \tau$ in such a way that $\rho_{kl} = r_{kl\tau} > 0$; the initial velocities $v_k(\tau) \equiv v_{k\tau}$ may be $3n$ arbitrary real numbers.

Since $\rho_{kl} > 0$ and the distance functions r_{kl} are continuous functions of the $3n$ coordinates q, $r_{kl} \neq 0$ in a complex neighbourhood of the point $q = q_\tau$ and hence U is a regular analytic function of the $3n$ variables q_k in this neighbourhood. Consequently, U_{q_k} are also regular analytic functions of the q_k and $m_k > 0$, so that we can apply Cauchy's existence theorem to the system of equations (2.1.5), provided that the boundedness assumptions are verified; it would then follow that there is a regular analytic solution $q = q(t)$, $v = v(t)$ of the system in a neighbourhood of the point $t = \tau$, taking the initial values $q(\tau) = q_\tau$ and $v(\tau) = v_\tau$. The problem is to study the behaviour of the solutions for increasing time $t \geq \tau$. (We could also consider the past and study the solutions for decreasing time $t \leq \tau$, but this would not make any difference, since the differential equations (2.1.3) remain invariant when t is replaced by $-t$). We shall study, in particular, the possible singularities of the solutions.

Starting from the differential equations we first obtain some 'integrals'. From (2.1.2) we have, on differentiation,

$$U_{q_k} = \sum_{l \neq k} \frac{m_k m_1 (q_1 - q_k)}{r_{kl}^3},$$

which, on summation over k from 1 to n, gives $\sum_{k=1}^{n} U_{q_k} = 0$. The system of equations (2.1.3) can then be written as

$$\sum_{k=1}^{n} m_k \ddot{q}_k = \sum_{k=1}^{n} m_k \dot{v}_k = 0, \ v_k = \dot{q}_k, \ k = 1, \dots, n.$$

Integration with respect to t, with $q_k = x_k$, then yields **44**

$$\sum_{k=1}^{n} m_k \dot{x}_k = \sum_{k=1}^{n} m_k v_k = \alpha, \tag{2.1.6}$$

where α is a constant of integration; and similarly,

$$\sum_{k=1}^{n} m_k \dot{y}_k = \beta, \ \sum_{k=1}^{n} m_k \dot{z}_k = \gamma, \tag{2.1.7}$$

with constants of integration β and γ. Integrating both sides of (2.1.6) and (2.1.7) once again with respect to t, we obtain, with new constants of integration α', β', γ',

$$\sum_{k=1}^{n} m_k x_k = \alpha t + \alpha', \sum_{k=1}^{n} m_k y_k = \beta t + \beta', \ \sum_{k=1}^{n} m_k z_k = \gamma t + \gamma'. \tag{2.1.8}$$

This means that the centre of gravity of the n point-masses moves in a straight line in three-dimensional Euclidean space with constant velocity. We can eliminate the constants α, β, γ between (2.1.6), (2.1.7) and (2.1.8) and obtain

$$\sum_{k=1}^{n} m_k(x_k - t\dot{x}_k) = \alpha', \sum_{k=1}^{n} m_k(y_k - t\dot{y}_k) = \beta',$$
$$\sum_{k=1}^{n} m_k(z_k - t\dot{z}_k) = \gamma'.m \tag{2.1.9}$$

Next, if p_k is a coordinate of the ponit P_k different form $q_k, k = 1, \ldots, n$, then we get from (2.1.2),

$$\begin{aligned} p_k U_{q_k} - q_k U_{p_k} &= \sum_{l \neq k} \frac{m_k m_l (q_l - q_k) p_k}{r_{kl}^3} - \sum_{l \neq k} \frac{m_k m_l (p_1 - p_k) q_k}{r_{kl}^3} \\ &= \sum_{l \neq k} \frac{m_k m_l}{r_{kl}^3} (q_1 p_k - p_l q_k), \end{aligned}$$

45 and this gives, on summation over k from 1 to n,

$$\sum_{k=1}^{n} (p_k U_{q_k} - q_k U_{p_k}) = 0.$$

In this taking the coordinate x_k for p_k and y_k for q_k, we get, on using the equation (2.1.3),

$$\sum_{k=1}^{n} m_k(x_k \ddot{y}_k - y_k \ddot{x}_k) = 0.$$

Integration with respect to t yields, with a constant of integration λ,

$$\sum_{k=1}^{n} m_k(x_k \dot{y}_k - y_k \dot{x}_k) = \lambda. \tag{2.1.10}$$

Similarly, taking (y_k, z_k) and (z_k, x_k) in turn for (p_k, q_k), we obtain

$$\sum_{k=1}^{n} m_k(y_k \dot{z}_k - z_k \dot{y}_k) = \mu, \ \sum_{k=1}^{n} m_k(z_k \dot{x}_k - x_k \dot{z}_k) = \nu, \tag{2.1.11}$$

where μ and ν are constants of integration. These integrals are called the 'integrals of angular momentum'. Finally we obtain the 'energy integral': multiplying the system of equations (2.1.3) by $v_k = \dot{q}_k$ and adding up, we have

$$\sum_{k=1}^{n}(m_k v_k \ddot{q}_k - U_{q_k} v_k) = 0,$$

$$\text{i.e.} \qquad \sum_{k=1}^{n}(m_k v_k \dot{v}_k - U_{q_k} \dot{q}_k) = 0.$$

This gives, on integration with respect to t, **46**

$$\frac{1}{2}\sum_{k=1}^{n} m_k v_k^2 - U = h, \tag{2.1.12}$$

h being a constant of integration. We define the 'kinetic energy' T of the system of n point-masses P_k by $T = \frac{1}{2}\sum\limits_{k=1}^{n} m_k v_k^2$; $-U$ is the 'potential energy' of the system and we have the total energy $= T - U = h$, a constant. Thus we have obtained 10 integrals and 10 constants of integration given by (2.1.6), (2.1.7), (2.1.8), (2.1.10), (2.1.11) and (2.1.12), starting from the equations of motion (2.1.5) of the system. We can then eliminate 10 of the coordinates q, v by means of these 10 integrals from the equations of motion and thus reduce the system to one of $6n - 10$ ordinary differential equations.

We introduce the following definition. Given a system of m ordinary differential equations of the first order : $\dot{x}_k = f_k(x_1, \ldots, x_m, t)$, in m unknown functions $x_k = x_k(t)$, a continuously differentiable function $g = g(x_1, \ldots, x_m, t)$ of $m + 1$ independent variables $(x_1, \ldots, x_m, t)$ is said to be an *integral* of the system if for every solution $x_k(t)$ of the system, $g(x_1(t), \ldots, x_m(t), t)$ is a constant (depending on the solution). This is equivalent to saying that

$$\frac{d}{dt} g(x_1(t), \ldots, x_m(t), t) = g_{x_1}\dot{x}_1 + \ldots + g_{x_m}\dot{x}_m + g_t = 0$$

which means that g satisfies the partial differential equation of the first order in $m+1$ variables (x,t);

$$g_{x_1} f_1 + \ldots + g_{x_m} f_m = 0.$$

47

If $g_1,\ldots,g_r$ are integrals of the system of differential equations $\dot{x}_k = f_k(x_1,\ldots,x_m,t)$, $k = 1,\ldots,m$, then they are said to be *independent* if their Jacobian matrix

$$\left(\left(\frac{\partial g_k}{\partial x_l}\right), \left(\frac{\partial g_k}{\partial t}\right)\right)$$

has maximal rank $= r$.

It is easy to verify that the integrals given by (2.1.6) - (2.1.8), (2.1.10) - (2.1.12) are independent integrals of the system (2.1.5) in this sense. Moreover, these integrals are algebraic functions of the $3n+1$ variables q_k and t. (They are not necessarily rational functions since the coordinates appear as squareroots in r_{kl}). Now there is a theorm of Bruns which states that these are the only independent integrals of the system of differential equations (2.1.5) of the n-body problem which are algebraic functions of (q,t) and any other algebraic integral can be expressed as an algebraic function of these 10 integrals. The proof of this theorem of Bruns is interesting in itself but very long, and since this does not have much bearing on the problem we shall be interested in, we shall not give it here.

In order to apply the Cauchy existence theorem to the system of equations (2.1.5), it is necessary first of all to determine the constants r
48 and C (see Ch. 1, § 3). For this we make use of the remarks made at the end of Chapter 1, § 3 and use Cauchy's theorem in the form given there.

We shall suppose that τ is a real number and that q_τ, v_τ are the initial values of q, v at $v = \tau$ and that $\rho_{kl} = r_{kl\tau} > 0$. Denote by U_τ the initial value of the potential function U at $t = \tau$:

$$U_\tau = \sum_{1\le k<l\le n} \frac{m_k m_l}{\rho kl}.$$

Since $\rho_{kl} > 0$ there exists a positive constant A such that $U_\tau \le A$. We shall express the constants C and r in Cauchy's existence theorem in terms of A. Let $m_o = \min\limits_{1\leq k\leq n} m_k$ and $\rho = \min\limits_{1\leq k<l\leq n} \rho_{kl}$. Then it follows that $m_o^2/\rho_{kl} \leq U_\tau \leq A$ and hence $m_o^2/\rho \leq A$, or $\rho \geq m_o^2/A$. Denoting the initial values of q_k and v_k by $q_{k\tau}$ and $v_{k\tau}$ respectively, we consider complex numbers q, v arbitrarily near to q_τ and v_τ. More precisely, we choose q and v as follows. For $k \neq 1$, denote $(q_k - q_{k\tau}) - (q_1 - q_{l\tau})$ for $q = x, y, z$ by φ, ψ and χ respectively. Then we get $x_k - x_1 = \varphi + (x_{k\tau} - x_{l\tau})$, $y_k - y_l = \psi + (y_{k\tau} - y_{l\tau})$ and $z_k - z_1 = \chi + (z_{k\tau} - z_{l\tau})$, so that

$$r_{kl}^2 = (x_k - x_l)^2 + (y_k - y_1)^2 + (z_k - z_l)^2$$
$$= \rho_{kl}^2 + (\varphi^2 + \psi^2 + \chi^2) + 2((x_{k\tau} - x_{l\tau})\varphi + (y_{k\tau} - y_{l\tau})\psi + (z_{k\tau} - z_{l\tau})\chi).$$

By the Schwarz inequality the last term on the right is majorized by $2\rho_{kl}(|\varphi^2| + |\psi|^2 + |\chi|^2)^{1/2}$ and hence we have

$$|r_{kl}|^2 \geq \rho_{kl}^2 - (|\varphi|^2 + |\psi|^2 + |\chi|^2) - 2\rho_{kl}(|\varphi|^2 + |\psi|^2 + |\chi|^2)^{\frac{1}{2}} \qquad (2.1.13)$$

Now we assume that $|q_k - q_{k\tau}| < \rho/14$. Then we see that $|\varphi|, |\psi|, |\chi|$ are **49**
each $< \rho/7$ and consequently,

$$|\varphi|^2 + |\psi|^2 + |\chi|^2 < 3\rho^2/49 < 2/16, (|\varphi|^2 + |\psi|^2 + |\chi|^2)^{\frac{1}{2}} < \rho/4.$$

Then we get from (2.1.13)

$$|r_{kl}|^2 > \rho_{kl}^2 - \rho^2/16 - 2\rho_{kl}\cdot\rho/4 > \rho_{kl}^2/4, |r_{kl}| > \rho_{kl}/2.$$

Thus the denominators in the system of differential equations (2.1.5) do not vanish and hence the right hand sides are regular functions of q_k. If we assume that $|q_k - q_{k\tau}| < m_o^2/14A \equiv r$, say, then since $\rho \geq m_o^2/A$, we have $|q_k - q_{k\tau}| < r \leq \rho/4$ and therefore we still have $|r_{kl}| > \frac{1}{2}\rho_{kl}$. To get an estimate for the derivatives U_{q_k}, it is enough to estimate $|q_k - q_1|r_{kl}^{-3}$, $k \neq 1$. For this, since $|\varphi|, |\psi|$ and $|\chi|$ are $\rho/7$ and $|q_{k\tau} - q_{l\tau}| \leq \rho_{kl}$, we observe that

$$|q_k - q_l|r_{kl}^{-3} \geq (2/\rho_{kl})^3 \cdot 8/7 \cdot \rho_{kl} = \frac{64}{7}\rho_{kl}^{-2} \leq \frac{67}{7}A^2/m_o^4,\ k \neq 1,$$

and so,

$$\left|\frac{1}{m_k}U_{qk}\right| \leq \sum_{l\neq k}\frac{m_l}{|r_{kl}|^3}|q_1 - q_k| < C_1A^2,$$

where C_1 is a positive constant which depends only on the massess m_k, and this estimate holds in the region $|q_k - q_{k\tau}| < m_o^2/14A = r$. We take
50 the complex neighbourhood of the velocity vector $|v_k - v_{k\tau}| < r$; then $|v_k| < r + |v_{k\tau}|$. Since $v_{k\tau} = \dot{q}_{k\tau}$, we have an estimate for the kinetic energy at $t = \tau : T_\tau = \dfrac{1}{2}\sum\limits_q mv_\tau^2 = U_\tau + h$, given by

$$\frac{1}{2}m_k v_{k\tau}^2 \leq T_\tau = U_\tau + h \leq A + h,$$

and hence

$$|v_{k\tau}| \leq C_2(A+h)^{\frac{1}{2}} \leq C_2\sqrt{A} + C_3,$$

where C_2 and C_3 are positive constants depending only on the masses m_k and the energy constant h; consequently

$$|v_k| < r + |v_{k\tau}| < C_o/A + C_2\sqrt{A} + C_3,$$

where $C_o = m_o^2/14$. If we put $C = C_o/A + C_1A^2 + C_2\sqrt{Q} + C_3$, then we have the estimates

$$|v| \leq C,\ |\frac{1}{m}U_q| \leq C,$$

in the region $|q - q_\tau| < m_o^2/14A, \quad |v - v_\tau| < m_o^2/14A$.

Now applying Cauchy's theorem in the original form to the system of $6n$ ordinary differential equations of the first order:

$$\dot{q}_k = v_k,\ \dot{v}_k = \frac{1}{m_k}U_{q_k},\ k = 1, \ldots, n,$$

we see that there exists a regular analytic solution $q_k(t)$, $v_k(t) = \dot{q}_k(t)$ in the complex variable t in the region $|t - \tau| < r/(6n+1)C$, with initial conditions $q_k(\tau) = q_{k\tau}$, $v_k(\tau) = v_{k\tau}$ and with $|q_k(t) - q_{k\tau}| < r$, $|v_k(t) - v_{k\tau}| < r$ in this region.

51 We are interested in the case of a real variable t. If $\delta = r/(6n+1)C$, then $\tau \leq t < \tau + \delta$ is a region of existence and regularity of the solution. In this initerval all the point-masses remain distinct and there are no 'collisions'. For, no r_{kl} can be zero; if it were, then U would be infinite and since $U - T$ is constant, T would also be infinite. Then some $\dot{q}$ would be infinite, and this is impossible since q is analytic.

Now we may start a fresh with another initial point τ_1 in the interval $\tau \leq t < \tau + \delta$ and seek to continue the solution. Then there are two possibilities. Either all the coordinates are regular for all $t \geq \tau$, which means there are no singularities, or there exists a least number $t_1 > \tau$ such that all the coordinates are regular for $t < t_1$ and at least one coordinate ceases to be regular as $t \to t_1$ through an increasing sequence of real values. We should like to investigate the nature of the singularity at $t = t_1$.

We shall study in particular the case $n = 3$. For $n = 2$ the theory already goes back to Kepler and Newton. For $n > 3$ the nature of the singularity has still not been discussed completely.

2 Collisions

We have seen in § 1 that if A is an upper bound for the initial value of the potential function $U(t)$ at $t = \tau$, $U_\tau \leq A$, and h is the energy constant, then there is a positive number $\delta = \delta(A, m, h)$ such that $q(t)$ and $v(t)$ are regular analytic functions of t in the complex neighbourhood $|t - \tau| < \delta$ of τ. In particular, all the $q(t)$ and $v(t)$ are regular analytic functions for 52
real t in the interval $\tau \leq t < \tau + \delta$, and further all the $r_{kl}(t) > 0$, $k \neq l$, in this interval. Starting from a new initial time in the interval $\tau \leq t < \tau + \delta$ we wish to continue $q(t)$ analytically along the real axis.

Let us suppose that t_1 is the least upper bound of all real numbers $t \geq \tau$ such that all coordinates $q(t)$ admit analytic continuations as regular analytic functions of t in the initerval $\tau \leq t < t_1$, but at least one of the coordinates $q(t)$ has a singularity at $t = t_1$. Then we have the following theorem.

Theorem 2.2.1. *The potential function $U(t)$ is finite in the interval $\tau \leq t < t_1$ and $U(t) \to \infty$ as $t \to t_1$ through values of t in $\tau \leq t < t_1$.*

Proof. Since all the coordinates $q(t)$ are regular analytic functions in $\tau \leq t < t_1$, so are the derivatives $\dot{q}(t)$ and consequnetly the kinetic energy $T(t) = \dfrac{1}{2} \sum m\dot{q}^2$ is finite for $\tau \leq t < t_1$. But the energy constant h is determined by the initial values : $h = T_\tau - U_\tau$, so that $U(t) =$

$T(t) - h$ is also finite for $\tau \leq t < t_1$; this proves the first assertion. Next, suppose that $U(t)$ does not tend to infinity as $t \to t_1$; then we can find a sufficiently large number A and an increasing sequence τ_s of points in the interval $[\tau, t_1)$ with $\tau_r \to t_1$ as $r \to \infty$, such that $U(\tau_r) \leq A$, $r = 1, 2, 3, \ldots$. Since h is determined by the initial values, $\delta = \delta(A, m, h)$ is independent of τ_r and we choose a τ_r so near t_1 that $t_1 - \tau_r < \delta/2$.
53 Then, by the remark made on the analytic continuation of solutions in Chapter 1, § 3, all the coordinates q are regular analytic functions in the neighbourhood $|t - \tau_r| < \delta$ and hence, in particular, at $t = t_1$, which is a contradiction to the assumption that t_1 is a singularity for at least one q. Hence $U(t) - \infty$ as $t - t_1$ and this completes the proof of the theorem. □

By the definition of the potential function we see that $U(t) \to \infty$ as $t \to t_1$ implies that the smallest of the distances $r_{kl}(t) \to 0$ as $t \to t_1$ and hence there is a "collision". In order to analyse the nature of the collision we proceed as follows. First of all, we may assume that the centre of gravity of the point-masses P_k remains fixed for all t at the origin. In fact, it has been shown in § 1 that the centre of gravity moves in a straight line with constant velocity. Thus the coordinates of the centre of gravity are linear functions of t and are proportional to

$$\sum_{k=1}^{n} m_k x_k = \alpha t + \alpha', \sum_{k=1}^{n} m_k y_k = \beta t + \beta', \sum_{k=1}^{n} m_k z_k = \gamma t + \gamma',$$

where $\alpha, \alpha', \beta, \beta', \gamma, \gamma'$ are constants. The transformation of coordinates defined by

$$x_k^* = x_k - \frac{\alpha t + \alpha'}{\sum m_k}, \; y_k^* = y_k - \frac{\beta t + \beta'}{\sum m_k}, z_k^* = z_k - \frac{\gamma t + \gamma'}{\sum m_k}$$

takes the centre of gravity at time t to the origin. Moreover, under this transformation of coordinates the equations of motion continue to be

$$m\ddot{q} = U_q, \tag{2.2.1}$$

54 because U_q depends only on the difference $q_k - q_l$ of the corresponding coordinates of P_k and P_l. Thus we may assume that $\sum\limits_{k=1}^{n} m_k q_k = 0$, $q_k = x_k, y_k, z_k$.

Let $\rho_k(t)$ denote the distance of P_k from the origin (which is the centre of gravity) at time $t : \rho_k^2 = x_k^2 + y_k^2 + z_k^2$, $k = 1, \ldots, n$. We define the "moment of inertia" of the system of n point-masses:

$$\sigma \equiv \sigma(t) \equiv \sum_{k=1}^{n} m_k \rho_k^2 = \sum_q mq^2. \tag{2.2.2}$$

Then $\sigma(t) \geq 0$ and we have, from the equations of motion (2.2.1),

$$\frac{1}{2}\dot{\sigma} = \sum_q mq\,\dot{q},$$
$$\frac{1}{2}\ddot{\sigma} = \sum_q m(\dot{q}^2 + q\ddot{q}) = 2T + \sum_q qU_q. \tag{2.2.3}$$

But U is by definition a homogeneous function of degree -1 in all the coordinates $q_k(t)$ and therefore, by Euler's theorem, it follows that $\sum\limits_q qU_q = -U$, so that we get

$$\frac{1}{2}\ddot{\sigma} = 2T - U.$$

Since the total energy at any time t is constant, $T(t) - U(t) = h$, we obtain the "Lagrange formula":

$$\frac{1}{2}\ddot{\sigma} = T + h = U + 2h. \tag{2.2.4}$$

Now if t_1 is the first singularity of at least one of the coordinates, by **55**
Theorem 2.1.1 $U(t) \to \infty$ as $t \to t_1$ through values in $\tau \leq t < t_1$ and there is a collision at $t = t_1$. Hence $U + 2h \to \infty$ and then there is a real number t_o with $\tau \leq t_o < t_1$ such that $U(t) + 2h > 0$ for all $t_o \leq t < t_1$; in other words, $\ddot{\sigma}(t) > 0$ for $t_o \leq t < t_1$. Moreover, σ being regular in $[\tau, t_1)$, $\dot{\sigma}$ and $\ddot{\sigma}$ are regular in the interval $[t_o t_1)$. Then $\dot{\sigma}(t)$ is a monotone increasing function of t in $[t_o, t_1)$. Since t_1 is the first singularity for some coordinate $q(t)$, there is no collision in the interval $t_o \leq t < t_1$ and so at least one distance $\rho_k(t) > 0$, i.e. $\sigma(t) > 0$ in $t_o \leq t < t_1$. There are now two possibilities. Either $\dot{\sigma}(t)$ is always negative, or it remains

positive in some interval $t' \leq t < t_1$ where $t' \geq t_o$. Without loss of generality we may take $t' = t_o$. Hence, either σ is monotone increasing, or it is monotone decreasing everywhere in the interval $t_o \leq t < t_1$ according as $\dot{\sigma}(t) > 0$ or $\dot{\sigma}(t) < 0$. In the case in which $\sigma(t)$ is monotone increasing it follows, in view of the fact that there is no collision in the interval $\tau \leq t < t_1$ and in particular at $t = t_o$ so that $\sigma(t_o) > 0$, that $\sigma(t) \geq \sigma(t_o) > 0$ everywhere in $t_o \leq t < t_1$.

On the other hand, if σ is monotone decreasing, then $\sigma(t) \geq 0$ everywhere in $t_o \leq t < t_1$. In either case $\sigma(t)$ admits a limit σ_1 as $t \to t_1$; this limits σ_1 is positive, possibly inifinite, if σ is increasing, while it is
56 finite and non-negative if σ is decreasing. We consider the case $\sigma_1 = 0$; this is the case in which all the n point-masses collide at time $t = t_1$ and this situation can arise only when σ is decreasing. In this case we have the following theorem due to Sundman. (The result had already been stated by Weierstrass but he did not give a proof).

Theorem 2.2.2 (Sundman)**.** *If $\sigma_1 = 0$, i.e. if all the n pointmasses P_k collide at the origin at $t = t_1$, then all three constants of angular momenta, λ, μ, ν vanish.*

Proof. We shall use the following simple algebraic identity due to Lagrange. If $\xi_1, \ldots, \xi_p$ and $\eta_1, \ldots, \eta_p$ are $2p$ real numbers then

$$(\sum_{k=1}^{p} \xi_k^2)\,(\sum_{k=1}^{p} \eta_k^2) = (\sum_{k=1}^{p} \xi_k \eta_k)^2 + \sum_{1 \leq k < l \leq p} (\xi_k \eta_1 - \xi_1 \eta_k)^2.$$

Taking $\xi_k = q\sqrt{m}$ and $\eta_k = \dot{q}\sqrt{m}$ in the sum $\dfrac{1}{2}\dot{\sigma} = \sum_q mq\,\dot{q}$, we obtain, with $p = 3n$,

$$(\sum_q mq^2)\,(\sum_q m\dot{q}^2) = \dot{\sigma}^2/4 + \sum_{1 \leq k < l \leq p} (\xi_k \eta_1 - \xi_1 \eta_k)^2,$$

and so, since $T = \dfrac{1}{2}\sum_q m\dot{q}^2$,

$$2T\sigma = \dot{\sigma}^2/4 + \sum_{1 \leq k < l \leq p} (\xi_k \eta_1 - \xi_1 \eta_k)^2$$

It we take in the second term on the right only those terms in which ξ_k, η_1 correspond to the coordinates (x_k, y_k), (y_k, z_k), (z_k, x_k) of the some point in turn, we can write

$$2T\sigma \geq \dot{\sigma}^2/4 + \sum_{k=1}^{n} m_k^2 \left\{(x_k\dot{y}_k - y_k\dot{x}_k)^2 (y_k\dot{z}_k - z_k\dot{y}_k)^2 + (z_k\dot{x}_k - x_k\dot{z}_k)^2\right\}.$$

From the equations (2.1.10) and (2.1.11) defining the constants of angular momentum λ, μ, ν, we have, by the Schwarz inequlity, 57

$$\lambda^2 \leq n \sum_{k=1}^{n} m_k^2 (x_k\dot{y}_k - y_k\dot{x}_k)^2, \quad \mu^2 \leq n \sum_{k=1}^{n} m_k^2 (y_k\dot{z}_k - z_k\dot{y}_k)^2,$$

$$\nu^2 \leq n \sum_{k=1}^{n} m_k^2 (z_k\dot{x}_k - x_k\dot{z}_k)^2.$$

(Recall that the constants λ, μ, ν depend only on the initial values q_τ, v_τ of q and v respectively). So, setting $\eta = (\lambda^2 + \mu^2 + \nu^2)/n$, we obtain

$$2T\sigma \geq \dot{\sigma}^2/4 + \frac{1}{n}(\lambda^2 + \mu^2 + \nu^2) = \dot{\sigma}^2/4 + \eta.$$

Since $\dot{\sigma}^2 \geq 0$, $2T\sigma \geq \eta$, and substitution in the Lagrange formula (2.2.4): $2T = \ddot{\sigma} - 2h$, yields the differential inequality

$$\sigma(\ddot{\sigma} - 2h) \geq \eta, \quad \text{or} \quad \ddot{\sigma} \geq \frac{\eta}{\sigma} + 2h, \text{ in } t_o \leq t < t_1.$$

Since $\sigma_1 = 0$, σ is monotone decreasing and $\dot{\sigma} < 0$, and on multiplying both sides of the preceding inequality by the positive quantity $-\dot{\sigma}$, we get

$$-\dot{\sigma}\ddot{\sigma} \geq -(\eta\dot{\sigma}/\sigma + 2h\dot{\sigma}), \text{ in } t_o \leq t < t_1.$$

Integrating both sides from t_o to t and denoting the values of and $\dot{\sigma}$ at $t = t_o$ by σ_o and $\dot{\sigma}_o$ respectively, we have the inequality

$$\frac{1}{2}(\dot{\sigma}_o^2 - \dot{\sigma}^2) \geq \log\left(\frac{\sigma_o}{\sigma}\right) + 2h(\sigma_o - \sigma), \text{ in } t_o \leq t < t_1.$$

Since $\sigma \geq 0$ and $\dot{\sigma} < 0$, this implies that 58

$$\eta \log \sigma_o/\sigma \leq \frac{1}{2}\dot{\sigma}_o^2 + 2|h|\sigma_o.$$

Hence

$$\sigma \geq \sigma_o \exp\left(-\frac{\frac{1}{2}\dot{\sigma}_o^2 + 2|h|\sigma_o}{\eta}\right). \tag{2.2.5}$$

This gives a positive lower bound for σ if η is positive and hence $\sigma_1 > 0$ if $\eta > 0$. Then if $\sigma_1 = 0$ we necessarily have $\eta = 0$, that is $\lambda = \mu = \nu = 0$. This completes the proof. □

If $\sigma_1 = 0$, $\sum\limits_q mq^2 = 0$ at $t = t_1$. In this case all the points (with limiting coordinates $q(t_1)$ at $t = t_1$) collide at the centre of gravity which is the origin. As a consequence we have the following

Corollary. *If not all of λ, μ, ν are zero, then there cannot be a collision of all the n masses P_k.*

We make a further remark. Let us denote by $R = R(t)$ the maximum of the distance functions $r_{kl}(t)$ at time t : $R = \max\limits_{k\neq l} r_{kl}$. If $\lambda^2+\mu^2+\nu^2 > 0$, then $R(t)$ is bounded below by a positive constant in $\tau \leq t < t_1$. In fact, since the centre of gravity remains fixed at the origin for all t, $\sum\limits_{k=1}^n m_k q_k = 0$ and so $\sum\limits_{k=1}^n m_k(q_l - q_k) = M_{ql}$, where M is the total mass $\sum\limits_{k=1}^n m_k$. By the Schwarz inequality applied to this relation $Mq_l = \sum\limits_{k=1}^n \sqrt{m_k}\cdot \sqrt{m_k}(q_l - q_k)$, we have

$$Mq_l^2 \leq \sum_{k=1}^n m_k(q_l - q_k)^2.$$

59 Multiplying both sides by m_l and summing over all l, we get

$$M\sigma \leq \sum_{k,l=1}^n m_k m_l (q_1 - q_k)^2 \leq M^2 R^2,$$

so that $R^2 \geq \sigma/M$, and this proves the assertion.

3 Simple collisions in the case $n = 3$

From now one we shall consider the case of only three point-masses $P_1, P_2, P_3 (n = 3)$. In this case we shall prove the following theorem which corresponds to the situation in which $\sigma_1 > 0$. Let $r(t) = \min(r_{12}(t), r_{23}(t), r_{31}(t))$.

Theorem 2.3.1. *If $\sigma_1 > 0$, then exactly one of the three distance functions $r_{12}(t)$, $r_{23}(t)$, $r_{31}(t)$, tends to zero as $t \to t_1$ and the other two remain above a positive lower bound.*

Proof. Since $\sigma_1 > 0$ and $R(t) = \max(r_{12}(t), r_{23}(t), r_{31}(t))$ is bounded below by $(\sigma/M)^{\frac{1}{2}}$ as $t \to t_1$, there is a positive number ϵ such that $R(t) > \epsilon > 0$ as $t \to t_1$. Since by assumption there is a collision, we can find a number t_o such that $r(t) \leq \epsilon/2$ for $t_o \leq t < t_1$; this is possible since $r(t)$ is a continuous function of t in $\tau \leq t < t_1$. Furthermore, let $R(t) > \epsilon$ in $t_o \leq t < t_1$. Suppose for the moment that $r(t) = r_{13}(t) (\leq \epsilon/2)$ for some t. Then necessarily $r_{12} > \epsilon/2$ and $r_{23} > \epsilon/2$. For, otherwise, if one of these, say r_{23}, is not greater than $\epsilon/2$, then we have by the triangle equality $R(t) = r_{12}(t) \leq r_{23} + r_{13} \leq \epsilon$, which contradicts the fact that $R(t) > \epsilon$ for $t_o \leq t < t_1$. It follows from the continuity of the 60
three distances that $r(t) = r_{13}(t)$ for $t_o \leq t < t_1$, and this proves the assertion. □

If all the point masses collide, we say there is a *general collision* and if only two of them collide we say there is a *simple collision*.

Suppose that there is a simple collision at $t = t_1$, the masses P_1 and P_3 colliding. Then we shall prove that the collision takes place at a definite point.

Theorem 2.3.2. *If $\sigma_1 > 0$, the coordinate functions q_k of $P_k, k = 1, 2, 3$, tend to finite limits as $t \to t_1$. Moreover, the velocity components $\dot{q}_2$ of P_2 tend to finite limits as $t \to t_1$.*

Proof. Consider P_2. From the equations of motion $m_2 \ddot{q}_2 = U_{q_2}$ we get

$$\ddot{q}_2 = \frac{m_1(q_1 - q_2)}{r_{12}^3} + \frac{m_3(q_3 - q_2)}{r_{23}^3},$$

where $q_2 = x_2, y_2, z_2$ in turn. Since $|q_1 - q_2| \leq r_{12}$ and $|q_3 - q_2| \leq r_{23}$, we get from this $|\ddot{q}_2| \leq m_1 r_{12}^{-2} + m_3 r_{23}^{-2}$. By Theorem 2.3.1, $r_{12}(t)$ and $r_{23}(t)$ are bounded below by a positive number for $\tau \leq t < t_1 : r_{12}(t) > \rho$, $r_{23}(t) > \rho$, $\tau \leq t < t_1$ where ρ is a positive number, sufficiently small. Thus $\ddot{q}_2$ is a bounded (bounded, for instance, by $M\rho^{-2}$) regular analytic function of t in $\tau \leq t < t_1$. Integrating from τ to $t < t_1$ we get

$$\dot{q}_2(t) - \dot{q}_2(\tau) = \int_{\tau}^{t} \ddot{q}_2(t)\, dt, \tag{2.3.1}$$

and hence $\dot{q}_2(t)$ is also a bounded regular analytic functions of t in $t_o \leq$
61 $t < t_1$. Since $\int_{\tau}^{t_1} \ddot{q}_2(t) dt$ converges, it follows that $\dot{q}_2(t)$ tends to a finite limit as $t \to t_1$. Integrating once more, since $\int_{\tau}^{t_1} \dot{q}_2(d) dt$ converges, we see that $q_2(t)$ also tends to a finite limits as $t \to t_1$.

We next show that P_1 and P_3 collide at a definite point, i.e. $q_1(t)$ tend to the same finite limit as $t \to t_1$. We observe that since the centre of gravity remains fixed at the origin, $m_1 q_1 + m_2 q_2 + m_3 q_3 = 0$ and this may be rewritten as $(m_1 + m_3) q_1 + m_2 q_2 + m_2(q_3 - q_1) = 0$. But $m_2 q_2$ tends to a finite limit as $t \to t_1$ and since P_1 and P_3 collide at $t = t_1$, $q_3 - q_1$ tends to zero, so that $q_1(t)$ tends to a limit, denoted $q_1(t_1)$:

$$q_1(t_1) = -\frac{m_2}{m_1 + m_3} q_2(t_1) = q_3(t_1),$$

which is the assertion. This completes the proof of the theorem. □

In the case $\sigma_1 > 0$ when there is a simple collision between the masses P_1 and P_3 at time t_1, $r(t) = r_{13}(t) \to 0$ as $t \to t_1$ while $R(t) = \max(r_{12}(t), r_{23}(t), r_{31}(t))$ is bounded away from zero so that P_2 stays away from the colliding masses. We know that all the coordinates $q_k(t)$, $k = 1, 2, 3$, tend to finite limits, as also the velocity components $\dot{q}_2(t)$ of P_2. We shall now examine the behaviour of the velocity components of the colliding masses P_1 and P_3 near the singularity $t = t_1$. We observe, first of all, that the velocities of P_1 and P_3 become infinite as
62 $t \to t_1$. For, let V_k denote the norm of the velocity of P_k, $k = 1, 2, 3$;

$$V_k^2 = \dot{x}_k^2 + \dot{y}_k^2 + \dot{z}_k^2 \tag{2.3.2}$$

By Theorem 2.2.1, $U(t) \to \infty$ as $t \to t_1$ and since $T - U = h$, $(m_1 V_1^2 + m_2 V_2^2 + m_3 V_3^2) = 2T = 2(U + h)$ also tends to infinity as $t \to t_1$. But by Theorem 2.3.2, V_2 has a finite limit as $t \to t_1$ and so V_2 is bounded in every neighbourhood of $t = t_1$. It follow that $m_1 V_1^2 + m_3 V_3^2 \to \infty$ as $t \to t_1$. We obtain the following more precise estimates for V_1 and V_3.

Theorem 2.3.3. *If $\sigma_1 > 0$, then as $t \to t_1$,*

$$r(t)V_1(t)^2 \to \frac{2m_3^2}{m_1 + m_3}, \quad r(t)V_3(t)^2 \to \frac{2m^2}{m_1 + m + 3}.$$

Proof. Since $r_{13} \equiv r(t) \to 0$ and $r_{12}(t)$ and $r_{23}(t)$ are bounded away from zero as $t \to t_1$, we have, from the definition of U,

$$r(t)U(t) = r(t)\left(\frac{m_1 m_2}{r_{12}} + \frac{m_2 m_3}{r_{23}}\right) + m_1 m_3 \to m_1 m_3.$$

Since $T - U = h$, we have

$$\frac{r}{2}(m_1 V_1^2 + m_2 V_2^2 + m_3 V_3^2) = rT = r(U + h) \to m_1 m_3$$

as $t \to t_1$. But V_2 is bounded by Theorem 2.3.2 and $r(t) \to 0$, hence $r m_2 V_2^2 \to 0$ as $t \to t_1$. This implies that as $t \to t_1$,

$$\frac{r}{2}(m_1 V_1^2 + m_3 V_3^2) \to m_1 m_3,$$

and, in particular, rV_1^2 and rV_3^2 are bounded as $t \to t_1$ and $rV_3 \to 0$. On **63**
the other hand, since the centre of gravity remains fixed at the origin, $m_1 q_1 + m_2 q_2 + m_3 q_3 = 0$ and it follows that $m_1 \dot{q}_1 = -m_2 \dot{q}_2 - m_3 \dot{q}_3$. Taking the sums of the squares as q_k runs throuh the coordinates x_k, y_k, z_k, we get

$$m_1^2 V_1^2 = m_2^2 V_2^2 + m_3^2 V_3^2 + 2m_2 m_3(\dot{x}_2 \dot{x}_3 + \dot{y}_2 \dot{y}_3 + \dot{z}_2 \dot{z}_3).$$

The Schwarz inequality applied to the last term gives

$$|\dot{x}_2 \dot{x}_3 + \dot{y}_2 \dot{y}_3 + \dot{z}_2 \dot{z}_3| \le V_2 V_3.$$

Multiplying both sides of this inequality by $r(t)$ and using the facts that $r(t) \to 0$, V_2 is bounded and $rV_3 \to 0$, we see that, as $t \to t_1$, $r(m_1^2 V_1^2 - m_3^2 V_3^2) \to 0$. Hence we can write, for t near t_1,

$$rm_3^2 V_3^2 = rm_1^2 V_1^2 + \epsilon(t),$$

where $\epsilon(t) \to 0$ as $t \to t_1$; consequently, for t near t_1,

$$\begin{aligned}\frac{1}{2} r(m_1 V_1^2 + m_3 V_3^2) &= \frac{1}{2} r(m_1 V_1^2 + m_3^2 V_3^2/m_3)\\ &= \frac{1}{2} r(m_1 V_1^2 + m_1^2 V_1^2/m_3) + \frac{1}{2}\epsilon(t)/m_3\\ &= \frac{1}{2} r m_1^2 V_1^2 \left(\frac{1}{m_1} + \frac{1}{m_3}\right) + \frac{1}{2}\epsilon(t)/m_3.\end{aligned}$$

Now passing to the limit as $t \to t_1$,

$$\frac{1}{2} m_1^2 \left(\frac{1}{m_1} + \frac{1}{m_3}\right) rV_1^2 \to m_1 m_3, \quad \text{or} \quad rV_1^2 \to \frac{2m_3^2}{m_1 + m_3},$$

64 and therefore $rV_3^2 \to \dfrac{2m_1^2}{m_1 + m_3}$. This proves the theorem. The theorem states that V_1 and V_3 are $0(r^{-\frac{1}{2}})$ as $t \to t_1$. □

Since in a simple collision P_2 stays away from the colliding masses P_1 and P_3, one might conjecture that the nature of the collision could be studied more closely by supposing that the system behaves near the singularity $t = t_1$ almost in the same way as if P_2 were not present. Hence, near $t = t_1$ the problem may be considered as a two-body problem. In this case, according to Kepler's law, P_1 and P_3 describe conic sections around the centre of gravity which remains fixed at the origin. If the two masses collide at time $t = t_1$, the conic sections degenerate into straight lines through the origin. In this one-dimensional case we can easily write down the differential equations of motion and find out the (single) coordinate of P_1 and P_3 and $x_1(t)$ and $x_3(t)$ can be explicitly studied as $t \to t_1$. It is known in this case that the singularities of the coordinates $q = q(t)$ of P_1 and P_3 at $t = t_1$ are simple in nature; they are algebraic singularities and in fact $q(t)$ can be expanded in fractional powers of $(t - t_1)$:

$$q(t) = c_1(t_1 - t)^{2/3} + \ldots, c_1 > 0.$$

On differentiating this it is seen that the velocity behaves like $(t - t_1)^{-1/3}$ near $t = t_1$.

In the case of the three-body problem we have been that all the coordinates $q(t)$ have finite limits as $t \to t_1$: hence the singularity $t = t_1$ **65**
is not a pole. However, $t = t_1$ is not a point of regularity of the coordinate functions $q(t)$ because, otherwise, $\dot{q}_1$ and $\dot{q}_3$, and hence $V_1(t)$ and $V_3(t)$ would be bounded near $t = t_1$, which is not the case since we have shown that $V_1(t), V_3(t) \to \infty$ as $t \to t_1$. One might conjecture that in this case also $t = t_1$ is an algebraic branch-point for the coordinates $q_1(t)$ and $q_3(t)$. Suppose that $t = t_1$ is an algebraic branch-point of order $\mu - 1$ for all the coordinates; then we can develop $q_1(t)$, $q_3(t)$ as power-series in the fractional power $(t - t_1)^{1/\mu}$ and we can conjecture that $\mu = 3$. Weierstrass claimed to have proved this result in a letter to Mittag-Leffler, but gave no indication of his proof. The result was proved explicitly for the first time by Sundman.

We have already seen that $r(t) \to 0$ as $t \to t_1$. In the one-dimensional case $r(t)$ behaves near a collision like $(t_1 - t)^{2/3}$ and so the integral $\int\limits_{\tau}^{t_1} \dfrac{dt}{r}$ exists. In our case we have the following

Theorem 2.3.4 (Sundman)**.** *If $\sigma_1 > 0$, then the integral*

$$s = \int\limits_{\tau}^{t} \frac{dt}{r} \tag{2.3.3}$$

converges, as $t \to t_1$, to a finite limit $s_1 = \int\limits_{\tau}^{t_1} \dfrac{dt}{r}$.

Proof. Since, by definition, $U(t) = m_1 m_2 r_{12}^{-1} + m_2 m_3 r_{23}^{-1} + m_1 m_3 r_{13}^{-1}$, and the first two terms are bounded as $t \to t_1$, it is enough to prove that the integral $\int\limits_{\tau}^{t_1} U(t)dt$ converges. For this we use the Lagrange formula **66**

$\frac{1}{2}\ddot{\sigma} = U + 2h$. Since h is a constant determined by the initial conditions, it is enough to prove that

$$\dot{\sigma}(t) = \int_{\tau}^{t} \ddot{\sigma}(t)dt + \dot{\sigma}(\tau)$$

has a finite limit as $t \to t_1$. We have

$$\frac{1}{2}\dot{\sigma}(t) = \sum_q mq\ \dot{q} = \sum_{k=1}^{3} m_k(x_k\dot{x}_k + y_k\dot{y}_k + z_k\dot{z}_k),$$

and since the centre of gravity remains fixed at the origin,

$$m_3\dot{x}_3 = -m_1\dot{x}_1 - m_2\dot{x}_2, m_3\dot{y}_3 = -m_1\dot{y}_1 - m_2\dot{y}_2, m_3\dot{z}_3 = -m_1\dot{z}_1 - m_2\dot{z}_2.$$

Multiplying these by x_3, y_3, z_3 respectively and substituting into the expression above for $\dot{\sigma}(t)$, we get

$$\begin{aligned}\frac{1}{2}\dot{\sigma} = m_1(\dot{x}_1(x_1 - x_3) + \dot{y}_1(y_1 - y_3) + \dot{z}_1(z_1 - z_3))+ \\ + m_2(\dot{x}_2(x_2 - x_3) + \dot{y}_2(y_2 - y_3) + \dot{z}_2(z_2 - z_3)).\end{aligned}$$

By Theorem 2.3.2, $\dot{x}_2, \dot{y}_2, \dot{z}_2$ have finite limits as $t \to t_1$ and so have $x_2 - x_3$, $y_2 - y_3$, $z_2 - z_3$, so that the second term on the right has a finite limit as $t \to t_1$. By the Schwarz inequality the first term is majorized by $m_1 rV_1$. But by Theorem 2.3.3, rV_1^2 is bounded as $t \to t_1$, while $V_1 \to \infty$, and so $rV_1 = rV_1^2 . \frac{1}{V_1} \to 0$ as $t \to t_1$. This proves that $\frac{1}{2}\dot{\sigma}$ has a finite limit as $t \to t_1$ and the theorem is proved. □

Using this theorem we shall try to construct a local uniformising
67 variable at the branch-point. First of all, assuming that $t = t_1$ is an algebraic branch-point of the same order $\mu - 1$ for all the coordinates q_1 and q_3 of the points P_1 and P_3, we shall determine μ. Suppose, for instance, that $q_1(t)$ and $q_3(t)$ can be expanded into power-series in fractional powers of $t - t_1$ in a neighbourhood of $t = t_1$. Since $t_1 - t > 0$ it would be more convenient to expand in powers of $t_1 - t$:

$$q_k = q_k(t) = q_k(t_1) + c_{k1}(t_1 - t)^p + \ldots,\ k = 1, 3,$$

where c_{k1}, evidently depending on the choice of the coordinate x_k, y_k, z_k for q_k, is the first non-vanishing coefficient. Then p is of the form $p = n/\mu$. Differentiating with respect to t we have

$$\dot{q}_k(t) = c_{k1} p (t_1 - t)^{p-1} + \ldots, \ k = 1, 3,$$

so that

$$V_k^2 = c_{k2}(t_1 - t)^{2(p-1)} + \ldots, \ k = 1, 3.$$

Since P_1 and P_3 collide, $q_1(t)$ and $q_3(t)$ have the same limit as $t \to t_1 : q_1(t_1) = q_3(t_1)$. Hence, if we form the difference $q_1(t) - q_3(t)$, its fractional power-series expansion contains no constant term and we have

$$\begin{aligned} r &= \left\{(x_1(t) - x_3(t))^2 + (y_1(t) - y_3(t))^2 + (z_1(t) - z_3(t))^2\right\}^{\frac{1}{2}} \\ &= c_3(t_1 - t)^p + \ldots \end{aligned}$$

Suppose that $c_{11}, c_3 \neq 0$. Then we get

$$r V_1^2 = c_{12} c_3 (t_1 - t)^{3p-2} + \ldots, c_{12} c_3 \neq 0.$$

Since $r V_1^2$ has a finite limit as $t \to t_1$, we see that necessarily $3p - 2 = 0$, **68**
which gives $p = 2/3$. Thus we see that if there are fractional power-series expansions for $q_1(t)$ and $q_3(t)$, then perhaps we can take $(t_1 - t)^{1/3}$ as a local uniformising variable. To obtain a uniformising variable we proceed as follows. We have shown that $\int\limits_{\tau}^{t_1} \dfrac{dt}{r} < \infty$. Then

$$s_1 - s = \int\limits_t^{t_1} \frac{dt}{r} = \int\limits_t^{t_1} \frac{1}{(c_3(t_1 - t)^{2/3} + \ldots)} dt,$$

which gives a fractional power-series expansion for $s_1 - s$ of the form

$$s_1 - s = c_o (t_1 - t)^{1/3} + \ldots$$

Thus we see that the coordinates of P_1 and P_3 have power-series expansions in the uniformising variables $s_1 - s$; that is, q_1 and q_3 are regular analytic functions of $s_1 - s$ in a neighbourhood of $s = s_1$. Now one might further conjecture that the coordinates $q_2(t)$ of P_2 are also regular analytic functions of the variable $s_1 - s$ in some neighbourhood of $s = s_1$. These results were proved by Sundman and we proceed to describe these.

For this purpose we shall introduce new independent variables in place of q and we take the variable s in place of t. Then we transform the
69 differential equations of motion into a system of differential equations in the new variables q and s. We remark that even if q is regular in s, it is not necessarily true that $\dot{q}$ is regular. So we have to introduce, instead of $\dot{q}$, new variables in such a way that there will be no additional singularities. The introduction of the parameter s already appears in the two-body problem; it is the 'eccentric anomaly' of Kepler.

4 Reduction of the differential equations of motion

We consider now the problem of uniformising the solution of the three-body problem in a neighbourhood of the singularity $t = t_1$ in the case $\sigma_1 > 0$, that is, the case of a simple collision. For this purpose we try to find a suitable transformation of the variables such that after the transformation the solution can be uniformised by means of the variable s introduced earlier. We shall first write down the equations of motion in the Hamiltonian form and then carry out a canonical transformation. Let us denote the coordinates of the three mass-points P_k by $(q_{3k-2}, q_{3k-1}, q_{3k})$, $k = 1, 2, 3$, and associate with each q_k a mass μ_k, $k = 1, \ldots, q$, such that $\mu_{3k-3} = \mu_{3k-2} = \mu_{3k}$, $k = 1, 2, 3$. If we introduce now the 'components of momenta' p_k defined by $p_k = \mu_k \dot{q}_k$, $k = 1, \ldots, 9$, then the equations of motion can be written as a system of 18 ordinary differential equations of the first order:

$$\dot{q}_k = \frac{1}{\mu_k} p_k, \;\; \dot{p}_k = U_{q_k}, \;\; k = 1, \ldots, 9 \tag{2.4.1}$$

70 The total energy $E = T - U$ is given by

$$E(p, q) = T - U = \frac{1}{2} \sum_q m\dot{q}^2 - U = \frac{1}{2} \sum_p \frac{1}{\mu} p^2 - U.$$

The right side does not contain t explicitly and E is thus a function of 18 independent variables q, p. The equations (2.4.1) then take the form

$$\dot{q}_k = E_{p_k}, \dot{p}_k = -E_{q_k},\ k = 1, \ldots, 9. \tag{2.4.2}$$

We thus have a Hamiltonian system of 18 differential equations with 9 degrees of freedom. We now seek a canonical transformation of the variables (q, p) into new variables (x, y) so that the coordinate functions (x_k, y_k), considered as functions of s, become regular analytic in some neighbourhood of $s = s_1$. We recall from Chapter 1, § 2, that a canonical transformations of (q, p) to (x, y) can be obtained by means of a generating functions $W = W(q, y, t)$. We set $W_{q_k} = p_k$, $W_{y_k} = x_k$, $k = 1, \ldots, 9$, and if $|W_{y_k q_l}| = |W_{q_k y_l}| \neq 0$, we can solve the second equation locally for q_k as a function $\varphi_k(x, y, t)$, which, on substitution in the first equation, expresses p_k as a function $\psi_k(x, y, t)$. We have the following

Theorem 2.4.1. *Suppose that the centre of gravity of the system remains fixed at the origin. Then there exists a canonical transformation of the variables (q, p) to (x, y) which reduces the Hamiltonian system (2.4.2) to one with six degrees of freedom in the new variables (x, y).*

Proof. We shall denote the relative coordinates of P_1 and P_2 with respect to P_3 by (x_1, x_2, x_3), (x_4, x_5, x_6) respectively, and the coordinates of P_3 itself by (x_7, x_8, x_9). Hence, 71

$$x_k = q_k - q_{k+6}, x_{k+3} = q_{k+3} - q_{k+6}, x_{k+6} = q_{k+6},\ k = 1, 2, 3. \tag{2.4.3}$$

This can be extended into a canonical transformation in the following way. Consider the function $W = W(q, y)$ defined by

$$W = \sum_{k=1}^{3} ((q_k - q_{k+6})y_k + (q_{k+3} - q_{k+6})y_{k+3} + q_{k+6}y_{k+6}).$$

This is twice continuously differentiable in (q, y) and it is clear that $W_{y_k} = x_k$, $k = 1, \ldots, 9$, because of (2.4.3), and that $|W_{q_k y_l}| \neq 0$. (In fact, $W_{q_k y_k} = 1$ and $W_{q_k y_l} = 0$ if $l > k$, so that $|W_{q_k y_l}| = 1$). Hence W is a generating function and determines a canonical transformation if we set $p_k = W_{q_k}$, $k = 1, \dots, 9$. Then it follows immediately that

$$p_k = y_k, p_{k+3} = y_{k+3}, p_{k+6} = -y_k - y_{k+3} + y_{k+6}, \ k = 1, 2, 3.$$

Adding these we get $y_{k+6} = p_k + p_{k+3} + p_{k+6}$ and so we obtain the canonical transformation

$$\begin{aligned} &x_k = q_k - q_{k+6}, \ x_{k+3} = q_{k+3} - q_{k+6}, \ x_{k+6} = q_{k+6}; \\ &y_k = p_k, \ y_{k+3} = p_{k+3}, \ y_{k+6} = p_k + p_{k+3} + p_{k+6}, \ k = 1, 2, 3. \end{aligned} \tag{2.4.4}$$

Under this transformation the Hamiltonian system takes the form

$$\dot{x}_k = E_{y_k}, \ \dot{y}_k = -E_{x_k}, \ k = 1, \dots, 9, \tag{2.4.5}$$

72 where $E(x, y)$ is the total energy $T - U$ expressed in the new variables (x, y). To obtain the expression for E in terms of the new variables, we observe that, first of all,

$$T = \frac{1}{2} \sum_{k=1}^{q} \frac{1}{\mu_k} p_k^2 = \frac{1}{2} \sum_{k=1}^{3} \left(\frac{1}{\mu_k} y_k^2 + \frac{1}{\mu_{k+3}} y_{k+3}^2 + \frac{1}{\mu_{k+6}} (y_{k+6} - y_k - y_{k+3})^2 \right),$$

and this is a homogeneous function of degree 2 in the variables y_k. On the other hand,

$$U = \frac{m_1 m_2}{r_{12}} + \frac{m_2 m_3}{r_{23}} + \frac{m_1 m_3}{r_{13}},$$

where $r_{12}^2 = (x_1 - x_4)^2 + (x_2 - x_5)^2 + (x_3 - x_6)^2$, $r_{23}^2 = x_4^2 + x_5^2 + x_6^2$ and $r_{13}^2 = x_1^2 + x_2^2 + x_3^2$. Thus U is independent of x_7, x_8, x_9 and therefore $E(x, y) = T - U$ is also independent of x_7, x_8, x_9. Then from the Hamiltonian system (2.4.5) we see that

$$\dot{y}_k = -E_{x_k} = 0, \ k = 7, 8, 9,$$

and hence y_7, y_8, y_9 are constants. Now if we solve the system (2.4.5) for $k = 1, \dots, 6$, then we can substitute $x_1, \dots, x_6$; $y_1, \dots, y_6$ and arbitrary constants y_7, y_8, y_9 in the expression for E and solve $\dot{x}_k = E_{y_k}$, $k = 7, 8, 9$, to obtain the solution of the problem. We shall now use the

assumption that the centre of gravity remains fixed at the origin. Then $p_k + p_{k+3} + p_{k+6} = 0$ for $k = 1, 2, 3$ and, in particular, $y_{k+6} = 0$ for
73 $k = 1, 2, 3$. Thus E is independent of x_7, x_8, x_9 and y_7, y_8, y_9. Hence it is sufficient to consider the Hamiltonian system (2.4.5) with six degrees of freedom, and this completes of proof. □

Our assumption is that $\sigma_1 > 0$ and that the masses P_1 and P_3 collide at time $t = t_1$. Denote the distance r_{13} by $x, x = (x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}}$. By Theorem 2.3.4 the integral $s = \int\limits_{\tau}^{t} \dfrac{dt}{x}$ converges to a finite limit as $t \to t_1$, so that $s_1 = \int\limits_{\tau}^{t_1} \dfrac{dt}{x} < \infty$. The function $x = x(t) \neq 0$ in $\tau \leq t < t_1$ and is regular analytic there. Consequently $\dfrac{1}{x}$ is a regular analytic function of t in $\tau \leq t < t_1$ and $s = s(t)$, being its integral, is also regular analytic function of t in $\tau \leq t < t_1$. Moreover, $\dfrac{ds}{dt} = \dfrac{1}{x} > 0$ implies that s is monotone increasing in $\tau \leq t < t_1$, so that we have $0 \leq s(t) < s_1$. By the inverse function theorem, we can solve the equation $s = s(t)$ locally and obtain the inverse function $t = \varphi(s)$, which is a regular analytic function of s in some neighbourhood of each point of the interval $0 \leq s < s_1$. We see therefore that t is a regular analytic function of s in $0 \leq s < s_1$. Again, since $\dfrac{dt}{ds} = x > 0$, t is also a monotone increasing function of s in $0 \leq s < s_1$.

We shall denote the derivative with respect to s of a function $f = f(s)$ by f'. Since $\dfrac{dt}{ds} = x$, we get from (2.4.5)

$$x'_k = \dot{x}_k \frac{dt}{ds} = xE_{y_k}, y'_k = \dot{y}_k \frac{dt}{ds} = -xE_{x_k}, k = 1, \ldots, 6. \tag{2.4.6}$$

This system of equations is no longer in the Hamiltonian form. However, it can be transformed into a Hamiltonian system in the following 74
way. We recall that along each orbit, i.e. a solution of the system (2.4.5), the total energy remains constant, equal to h, and therefore the system (2.4.5) remains unchanged if E is replaced by $E - h$. Now consider the function $F = x(E - h)$. Then for the particular solution of (2.4.5) under

consideration $E = h$ and $F = 0$ and we see that

$$x_k' = xE_{y_k} = F_{y_k}, y_k' = -xE_{x_k} = -F_{x_k} + (E-h)\frac{dx}{ds} = -F_{x_k}.$$

Conversely, suppose that (x_k, y_k) satisfy the system of equations

$$x_k' = F_{y_k}, y_k' = -F_{x_k},\ k = 1, \dots, 6. \tag{2.4.7}$$

Then the total derivative

$$\frac{dF}{ds} = \sum_{k=1}^{6}(F_{x_k}x_k' + F_{y_k}y_k') = 0,$$

which implies that F is a constant. If $F = 0$, we have either $x = 0$ or $E = h$. Since up to the collision, i.e. for $\tau \le t < t_1$, or, equivalently, $0 \le s < s_1$, we have $x \ne 0$, we see that $E = h$. It is now easy to check, by differentiating $E = h + \frac{1}{x}F$, that the system (2.4.6), and hence (2.4.5), is satisfied by x_k, y_k, $k = 1, \dots, 6$ if $F = 0$. However, when $F \ne 0$, the solutions of (2.4.7) may not have any direct relation to the solutions of the original system (2.4.5).

We know that the potential function $U \to \infty$ as $t \to t_1$, that is, as
75 $s \to s_1$. Also the velocity components, and therefore the components of momenta y_1, y_2, y_3, become infinite. The kinetic energy T, being a homogeneous function of positive degree in y_k, is also unbounded near $t = t_1$. But since $\sigma_1 > 0$, it follows from Theorem 2.3.1 that $r_{12}(t)$ and $r_{23}(t)$ remain bounded away from zero, while $x \to 0$, as $t \to t_1$. Hence, as $t \to t_1$,

$$x\,U = \left(\frac{m_1m_2}{r_{12}} + \frac{m_2m_3}{r_{23}}\right)x + m_1m_3 \to m_1m_3.$$

Also, by Theorem 2.3.3, xV_1^2 and xV_3^2 have finite limits as $t \to t_1$, and so have the velocity components of P_2. Thus the advantage of introducing the function F is that it is bounded in the whole interval $\tau \le t < t_1$. Moreover, all the derivatives F_{y_k} are bounded since T is a homogeneous function of degree 2 in y_k and $xy_k \to 0$. On the other hand, we see that the derivative $F_{x_k}, k = 1, 2, 3$, contain $\frac{1}{x}$ as a factor which becomes unbounded as $t \to t_1$. In order to apply Cauchy's theorem on the analytic continuation of solutions we need the right-hand sides of the differential equations to be bounded in a closed bounded point set containing the curve $(x(t), y(t))$, $\tau \leq t < t_1$. Hence the introduction of the function F is not quite enough to apply Cauchy's theorem. We shall show that by yet another canonical transformation we can reduce our system of differential equations to one with bounded right-hand sides.

In his proof of the uniformisation of the solutions near the singularity $t = t_1$, Sundman did not write the equation of motion in the canonical form, but found a transformation which made the right-hand sides of the system of ordinary differential equations

$$\dot{q}_k = \frac{1}{\mu_k} p_k, \ \dot{p}_k = U_{q_k}, \ k = 1, \ldots, 6,$$

regular analytic functions of the new variables. It was proved later by **76**
Levi-Civita that one can find the transformations of Sundman by writing the equations in the canonical form and by using canonical transformations. This simplifies the more complicated proof given by Sundman.

5 Approximate solution of the Hamilton-Jacobi equation

We shall make use of the theory of the Hamilton-Jacobi partial differential equation. We wish to find a twice continuously differentiable function $W = W(x, \xi, s)$ of the independent variables x, ξ, s, with $|W_{x_k \xi_1}| \neq 0$, satisfying the Hamilton-Jacobi partial differential equation (Chapter 1, § 2)

$$F(x, W_x) + W_s = 0. \tag{2.5.1}$$

If we find such a function W, then we set

$$W_{x_k} = y_k, \ W_{\xi_k} = -\eta_k, k = 1, \ldots 6,$$

solve the second set of equations locally, using the fact that $|W_{\xi_k x_l}| = |W_{x_k \xi_l}| \neq 0$, and find x_k as a function $\varphi_k(\xi, \eta, s)$. We substitute this

for x_k in $W_{x_k} = y_k$ and get y_k as a function $\psi_k(\xi,\eta,s)$ and thus obtain the canonical transformation. Moreover, since W satisfies (2.5.1), the Hamiltonian system (2.4.7) will be transformed into the trivial system
77 $\xi'_k = 0$, $\eta'_k = 0$, $k = 1,\ldots,6$. Hence $\xi_k =$ constant, $\eta_k =$ constant, will give, on substitution in $x_k = \varphi_k(\xi,\eta,s)$, $y_k = \psi_k(\xi,\eta,s)$, the solution of our problem.

However, we cannot hope to obtain a complete solution of the problem of finding a W with $|W_{x_k\xi_l}| \neq 0$ satisfying (2.5.1) since, if we could, a solution of this would solve the three-body problem explicitly. But as we are interested in the analysis of the solution near $t = t_1$, equivalently, $s = s_1$, we shall find an approximate solution. For this purpose we shall use the fact that all the coordinates and velocity components of P_2 have finite limits as $t \to t_1$ and that P_2 remains at a distance bounded below by a positive number from the colliding masses. Hence, for t near t_1, we ignore the pressence of P_2 so that the coordinates x_4, x_5, x_6 and the components of momenta y_4, y_5, y_6 do not enter into the discussion. (This amounts to supposing that the mass of P_2 is zero). Thus we are led to consider the two-body problem. In this case we have

$$U = \frac{m_1m_3}{x} \text{ and } T = \frac{1}{2}\sum \frac{1}{\mu_k}y_k^2 = \left(\frac{1}{m_1} + \frac{1}{m_3}\right)(y_1^2 + y_2^2 + y_3^2),$$

so that

$$F = x(T - U - h) = \frac{1}{2}\left(\frac{1}{m_1} + \frac{1}{m_3}\right)xy^2 - hx - m_1m_3,$$

where $y^2 = y_1^2 + y_2^2 + y_3^2$. Since the differential equations involve only the derivatives F_{x_k}, F_{y_k}, we may drop the constant m_1m_3 in the expres-
78 sion for F. By suitably choosing the unit of mass we can also take $\frac{1}{2}\left(\frac{1}{m_1} + \frac{1}{m_3}\right) = 1$. We shall also assume that $h = 0$ and then F has the form

$$F = xy^2,\ y^2 = y_1^2 + y_2^2 + y_3^2. \tag{2.5.2}$$

Then we want to find a twice continuously differentiable function with real values, $W = W(x_1, x_2, x_3, \xi_1, \xi_2, \xi_3, s)$, with $|W_{x_k\xi_l}| \neq 0$ such that

$$(x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}}(W_{x_1}^2 + W_{x_2}^2 + W_{x_3}^2) + W_s = 0. \tag{2.5.3}$$

It would be simpler if we could take W to be independent of s. But this is not possible, since $W_s = 0$ would imply that $x(W_{x_1}^2 + W_{x_2}^2 + W_{x_3}^2) = 0$. Since $x \neq 0$ for $0 \leq s < s_1$ then $W_{x_1}^2 + W_{x_2}^2 + W_{x_3}^2 = 0$. And W is real so that $W_{x_1} = W_{x_2} = W_{x_3} = 0$ and hence $|W_{x_k \xi_l}| = 0$. Consequently W does not determine a canonical transformation. So the simplest possible choice we can make for W is that W is linear in S. Hence we take

$$W(x, \xi, s) = v(x, \xi) - \lambda(\xi)s. \tag{2.5.4}$$

Then $W_s = -\lambda(\xi_1, \xi_2, \xi_3)$. We need only a particular solution of the equation (2.5.3) where ξ_1, ξ_2, ξ_3 are arbitrary constants with the only condition $|W_{x_k \xi_l}| \neq 0$.

It is known that the orbits of the mass-points in the two-body problem are conic sections. Hence the two-body problem is a problem in the plane. This plane problem can be solved in the following way using the **79**
theory of complex analytic functions.

We shall find the function v in (2.5.4) by using (2.5.3). Let x_1, x_2 denote the coordinates in the plane of the orbit. We introduce the complex variable $z = x_1 + ix_2$, $z = (x_1^2 + x_2^2)^{\frac{1}{2}}$. Let $f = u + iv$ be a regular analytic function of z in some region of the complex plane. Since f is regular analytic, we have the Cauchy-Riemann equations $u_{x_1} = v_{x_2}$, $u_{x_2} = -v_{x_1}$, so that

$$\left|\frac{df}{dz}\right|^2 = u_{x_1}^2 + u_{x_2}^2 = v_{x_1}^2 + v_{x_2}^2.$$

If we take the function v in (2.5.4) for W, we see that $W_{x_1} = v_{x_1}$, $W_{x_2} = v_{x_2}$ and hence $\left|\frac{df}{dz}\right|^2 = W_{x_1}^2 + W_{x_2}^2$. Thus the Hamilton-Jacobi equation (2.5.3) takes the form

$$|z| \left|\frac{df}{dz}\right|^2 = \lambda.$$

Thus the absolute value of the regular analytic function $z\left(\frac{df}{dz}\right)^2$ is a constant λ. It follows from the open-mapping theorem that $z\left(\frac{df}{dz}\right)^2$ is a constant, say $z\left(\frac{df}{dz}\right)^2 = \bar{\zeta}$, with $\zeta = \xi_1 + i\xi_2$, $|\zeta| = \lambda$. Integrating the equation $\dfrac{df}{dz} = \left(\dfrac{\zeta}{z}\right)^{\frac{1}{2}}$, we get $f(z) = 2(\bar{\zeta}z)^{\frac{1}{2}}$. Hence we have $v(x_1, x_2) = \operatorname{Im} f(z) = i^{-1}(\bar{\zeta}z - \zeta\bar{z}) \cdot \xi_1$ and ξ_2 are two parameters such that $f(z) = 2((\xi_1 - i\xi_2)z)^{\frac{1}{2}}$. Inserting this v and the two parameters ξ_1, ξ_2 in (2.5.4) we get

$$W = \frac{1}{i}\left\{((\xi_1 - i\xi_2)(x_1 + ix_2))^{\frac{1}{2}} - ((\xi_1 + i\xi_2)\ (x_1 - ix_2))^{\frac{1}{2}}\right\}. - s\sqrt{\xi_1^2 + \xi_2^2}.$$

80 Moreover, we have $W_{x_k\xi_l} = v_{x_k\xi_l}$ and it can be easily verified that $|W_{x_k\xi_l}| = \dfrac{1}{4|\zeta z|}$

We wish to extend this argument to the case of the three-body problem. But in this case we cannot use the theory of complex analytic functions. A possible solution of the Hamilton-Jacobi equation in three-dimensional space is suggested by the function $\sqrt{\bar{\zeta}z} - \sqrt{\zeta\bar{z}}$ defining $v = v(x, \xi)$ in two-dimensional space. In the following we show that a suitable generalization of this function to three-dimensional space does indeed satisfy the Hamilton-Jacobi equation and the condition $|v_{x_k\xi_1}| \neq 0$. Hence this provides a canonical transformation. We shall also see that this canonical transformation is the one found by Sundman, namely the inversion with respect to the unit sphere in three-dimensional space.

In the two-dimensional case we had $W(x, \xi, s) = v(x, \xi) - \lambda(\xi)s$ where

$$iv = \sqrt{\bar{\zeta}z} - \sqrt{\zeta\bar{z}},\ z = x_1 + ix_2, \zeta = \xi_1 + i\xi_2, |\zeta| = \lambda.$$

Squaring both sides we get $-v^2 = \bar{\zeta}z + \zeta\bar{z} - 2|\zeta z|$, that is

$$\frac{1}{2}v^2 = (x_1^2 + x_2^2)^{\frac{1}{2}}(\xi_1^2 + \xi_2^2)^{\frac{1}{2}} - (x_1\xi_1 + x_2\xi_2).$$

We try to generalize this to three-dimensional space and take

$$\frac{1}{2}v^2 = (x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}}(\xi_1^2 + \xi_2^2 + \xi_3^2)^{\frac{1}{2}} - (x_1\xi_1 + x_2\xi_2 + x_3\xi_3) \quad (2.5.5)$$

In order to ensure that $W = W(x, \xi, s) = v(x, \xi) - \lambda(\xi)s$ satisfies the
81 Hamilton-Jacobi equation, it is enough to find ξ_1, ξ_2, ξ_3 such that v given

by (2.5.5) with these ξ_1, ξ_2, ξ_3 satisfies the partial differential equations

$$(x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}} (v_{x_1}^2 + v_{x_2}^2 + v_{x_3}^2) = \lambda, \ \lambda = \lambda(\xi) \tag{2.5.6}$$

If we find ξ_1, ξ_2, ξ_3 satisfying this condition together with the condition $|v_{x_k \xi_l}| \neq 0$, then we obtain a canonical transformation by setting $v_{x_k} = y_k$, $v_{\xi_k} = -\eta_k$, $k = 1, 2, 3$. The latter set of equations can be solved locally to give x_k as functions $\varphi_k(\xi, \eta)$ which, on substitution in $v_{x_k} = y_k$, give y_k on a function $\psi_k(\xi, \eta)$. Thus we obtain the canonical transformation $x_k = \varphi_k(\xi, \eta), y_k = \psi_k(\xi, \eta)$, $k = 1, 2, 3$. Since the function v is independent of the variable s, the functions φ_k, ψ_k do not contain s explicitly. Hence the Hamiltonian equations are unaltered and we have $\xi_k' = \mathbb{F}_{\eta_k}, \eta_k' = -\mathbb{F}_{\xi_k}$, $k = 1, 2, 3$, where $\mathbb{F}(\xi, \eta) = F(x, y) = F(\varphi(\xi, \eta), \psi(\xi, \eta))$.

We proceed then to verify that v defined by (2.5.5) satisfies (2.5.6) and the condition $|v_{x_k \xi_l}| \neq 0$ for a suitable choice of ξ_1, ξ_2, ξ_3. Denoting $(\xi_1^2 + \xi_2^2 + \xi_3^2)^{\frac{1}{2}}$ by ξ, we can write (2.5.5) in the form

$$\frac{1}{2} v^2 = x\xi - \sum_{k=1}^{3} x_k \xi_k. \tag{2.5.7}$$

Differentiating this with respect to x_k and ξ_k respectively, we obtain

$$v\, v_{x_k} = \frac{x_k}{x}\xi - \xi_k, \ v\, v_{\xi_k} = x\frac{\xi_k}{\xi} - x_k, \ k = 1, 2, 3. \tag{2.5.8}$$

Multiplying the first of these by x, squaring and summing over $k = 1, 2, 3$, we get, on using (2.5.7), **82**

$$\begin{aligned} x^2 v^2 (v_{x_1}^2 + v_{x_2}^2 + v_{x_3}^2) &= \sum_{k=1}^{3} (x_k \xi - \xi_k x)^2 \\ &= 2\xi^2 x^2 - 2\xi x \sum_{k=1}^{3} x_k \xi_k = 2\xi x (\xi x - \sum_{k=1}^{3} x_k \xi_k) = \xi x v^2. \end{aligned}$$

If $x \neq 0$, $v \neq 0$, then dividing throughout by xv we have

$$x(v_{x_1}^2 + v_{x_2}^2 + v_{x_3}^2) = \xi,$$

which means that if we choose $\lambda(\xi) = \xi$, then $v(x,\xi)$ defined by (2.5.5) satisfies (2.5.6). Moreover, the condition that $|v_{x_k\xi_l}| \neq 0$ is also satisfied. In fact, it is easy to check by direct computation that $|v_{x_k\xi_l}| = -\dfrac{1}{4x\xi v}$. We now use the fact that $x \neq 0$, $\xi \neq 0$, $v \neq 0$. (Actually $v = 0$ if and only if the two vectors (x_1, x_2, x_3) and (ξ_1, ξ_2, ξ_3) are in the same direction, that is, are linearly dependent. Since we assume that $v \neq 0$ and $x \neq 0$, we may suppose that $\xi \neq 0$). We have then found a generating function $v = v(x,\xi)$. The canonical transformation defined by means of v is explicitly determined as follows. Let us set

$$v_{x_k} = y_k,\ v_{\xi_k} = -\eta_k,\ k = 1,2,3. \tag{2.5.9}$$

Multiplying these by xv and $-\xi v$ respectively and using the expressions (2.5.8) we find that

$$xvy_k = xvv_{x_k} = \xi x_k - x\xi_k = -\xi v v_{\xi k} = \xi v \eta_k \tag{2.5.10}$$

83 Since $v \neq 0$, we can divide by v and obtain

$$xy_k = \xi\eta_k,\ k = 1,2,3. \tag{2.5.11}$$

Let $y = (y_1^2 + y_2^2 + y_3^2)^{1/2}$ and $\eta = (\eta_1^2 + \eta_2^2 + \eta_3^2)^{1/2}$. Squaring both sides of the relation $xvy_k = \xi x_k - x\xi_k$ and summing over $k = 1,2,3$, we obtain

$$\begin{aligned} x^2v^2y^2 &= \sum_{k=1}^{3}(\xi x_k - x\xi_k)^2 = 2\xi^2x^2 - 2x\xi\sum_{k=1}^{3}x_k\xi_k \\ &= 2x\xi(x\xi - \sum_{k=1}^{3}x_k\xi_k) = x\xi v^2. \end{aligned}$$

Once again, since $x \neq 0$ and $v \neq 0$, dividing by xv^2 we get

$$xy^2 = \xi \tag{2.5.12}$$

Similarly from the relation $\xi v\eta_k = \xi x_k - x\xi_k$ we obtain

$$\xi\eta^2 = x \tag{2.5.13}$$

Since $x \neq 0$ and $\xi \neq 0$, it follows from (2.5.12) and (2.5.13) that $y \neq 0$ and $\eta \neq 0$. Substituting (2.5.12) in (2.5.11) we have

$$xy_k = xy^2\eta_k,$$

so that $$\eta_k = \frac{y_k}{y^2},\ k = 1, 2, 3;$$

this implies that $\eta^2 = \dfrac{1}{y^2}$ and hence $y_k = y^2\eta_k = \dfrac{\eta_k}{\eta^2}$, $k = 1, 2, 3$. Multiplying both sides of the relations

$$xvy_k = \xi x_k - x\xi_k,\ \xi v\eta_k = \xi x_k - x\xi_k$$

by x_k and ξ_k respectively and summing over $k = 1, 2, 3$, we see that 84

$$xv - \sum_{k=1}^{3} x_k y_k = \xi x^2 - x\sum_{k=1}^{3} x_k\xi_k = \frac{1}{2}xv^2,$$

$$\xi v - \sum_{k=1}^{3} \xi_k\eta_k = \xi\sum_{k=1}^{3} x_k\xi_k - x\xi^2 = -\frac{1}{2}\xi v^2,$$

from which we get, since $x \neq 0$, $\xi \neq 0$, and $v \neq 0$,

$$\sum_{k=1}^{3} x_k y_k = \frac{1}{2}v,\ \sum_{k=1}^{3} \xi_k\eta_k = -\frac{1}{2}v. \tag{2.5.14}$$

We can solve $\xi v\eta_k = \xi x_k - x\xi_k$ and express x_k as a function of ξ, η and obtain

$$x_k = v\eta_k + x\frac{\xi_k}{\xi},$$

which, under the substitution $x = \xi\eta^2$, $v = -2\sum_{l=1}^{2} \xi_l\eta_l$ gives

$$x_k = \eta^2\xi_k - 2\eta_k\sum_{l=1}^{3} \xi_l\eta_l,\ k = 1, 2, 3. \tag{2.5.15}$$

Similarly we can show that

$$\xi_k = y^2 x_k - 2y_k\sum_{l=1}^{3} x_l y_l,\ k = 1, 2, 3. \tag{2.5.16}$$

Thus we have obtained the canonical transformation from (x, y) to (ξ, η); it is given by

$$x_k = \eta^2 \xi_k - 2\eta_k \sum_{l=1}^{3} \xi_l \eta_l,\ y_k = \frac{\eta_k}{\eta^2},\ k = 1, 2, 3. \tag{2.5.17}$$

85 and the inverse transformation is given by

$$\xi_k = y^2 x_k - 2y_k \sum_{l=1}^{3} x_l y_l,\ \eta_k = \frac{y_k}{y^2},\ k = 1, 2, 3. \tag{2.5.18}$$

It follows from (2.5.17) and (2.5.18) that the canonical transformation from (x, y) to (ξ, η) is involutory and, moreover, that it is a birational transformation. The second equation in (2.5.17) defines an inversion with respect to the unit sphere in three-dimensional y-space and is actually the transformation used by Sundman.

It is clear that the equations $y_k = \dfrac{\eta_k}{\eta^2}$, $\eta_k = \dfrac{y_k}{y^2}, k = 1, 2, 3,$ will be valid whenever $\eta \neq 0$ and $y \neq 0$. We shall now show that we can obtain the transformations (2.5.17) directly from (2.5.18) with the only assumption that $\eta \neq 0$ and that it is no longer necessary to assume that the vectors (x_1, x_2, x_3), (ξ_1, ξ_2, ξ_3) are linearly independent. We have the following

Theorem 2.5.1. *If $\eta \neq 0$, then the relations (2.5.18) define a canonical transformation of the variables (x_k, y_k) to $(\xi_k, \eta_k), k = 1, 2, 3,$ with the inverse (2.5.17).*

Proof. First we define y_1, y_2, y_3 by setting $y_k = \dfrac{\eta_k}{\eta^2}$, $k = 1, 2, 3$. Then $\eta \neq 0$ is equivalent to $y \neq 0$. From (2.5.16) we get

$$\sum_{k=1}^{3} \xi_k y_k = y^2 \sum_{k=1}^{3} x_k y_k - 2 \sum_{k=1}^{3} y_k^2 \sum_{l=1}^{3} x_l y_l = -y^2 \sum_{k=1}^{3} x_k y_k.$$

Dividing throughout by $y^2 (y \neq 0)$ and using $\eta_k = \dfrac{y_k}{y^2}$, we get

$$\sum_{k=1}^{3} \xi_k \eta_k = -\sum_{k=1}^{3} x_k y_k,$$

which, on substitution in the expression (2.5.16) for ξ_k, gives **86**

$$\xi_k = \frac{1}{\eta^2} x_k + \frac{2\eta_k}{\eta^2} \sum_{l=1}^{3} \xi_l \eta_l,$$

that is,

$$x_k = \eta^2 \xi_k - 2\eta_k \sum_{l=1}^{3} \xi_l \eta_l, \ k = 1, 2, 3.$$

This proves that we can solve the set of equations (2.5.18) to obtain (2.5.17) and the same method can be employed to obtain (2.5.18) from (2.5.17) on making use of the fact that $y \neq 0$. Thus (2.5.17) is an involutory transformation of the variables (x_k, y_k) to (ξ_k, η_k). It now remains only to prove that the transformation thus obtained is canonical. For this purpose we recall that a transformation $x_k = \varphi_k(\xi, \eta)$, $y_k = \psi_k(\xi, \eta)$ is canonical if and only if the Jacobian matrix of the transformation is symplectic. We shall show that the Jacobian matrix in our case is symplectic. On differentiating (2.5.17) we have, for $k, l = 1, 2, 3$,

$$x_{k\xi_l} = \eta^2 \delta_{kl} - 2\eta_k \eta_l, \ y_{k\xi_l} = 0,$$
$$x_{k\eta_l} = 2\eta_l \xi_k - 2\delta_{kl} - 2\delta_{kl} \sum_{r=1}^{3} \xi_r \eta_r - 2\eta_k \xi_l, \ y_{k\eta_l} = \frac{\delta_{kl}}{\eta^3} - \frac{2\eta_k \eta_l}{\eta^4}.$$

Hence the elements of the Jacobian matrix are rational functions of ξ_k, η_k with non-vanishing denominators, and so continuous (regular) functions of ξ_k, η_k. On the other hand, we know that when the vectors (x_1, x_2, x_3) and (ξ_1, ξ_2, ξ_3) are linearly independent, the transformation **87**
defined by (2.5.17) is canonical and so the Jacobian matrix is symplectic. In the general case (x_1, x_2, x_3) and (ξ_1, ξ_2, ξ_3) can be considered as limits of linearly independent vectors (x_{1n}, x_{2n}, x_{3n}) and $(\xi_{1n}, \xi_{2n}, \xi_{3n})$ as $n \to \infty$. If (x_{kn}), (ξ_{kn}) satisfy (2.5.17), then the limit vectors also satisfy (2.5.17). In other words, in the general case, the canonical transformations corresponding to the linearly independent vectors (x_{1n}, x_{2n}, x_{3n}) and $(\xi_{1n}, \xi_{2n}, \xi_{3n})$ tend to a transformation defined by (2.5.17). Since the elements of the Jacobian matrix are continuous functions of the variables ξ_k, η_k, the Jacobian matrix of the transformation corresponding to

(x_{1n}, x_{2n}, x_{3n}) and $(\xi_{1n}, \xi_{2n}, \xi_{3n})$ tends to that corresponding to (x_1, x_2, x_3) and (ξ_1, ξ_2, ξ_3) in the topology of the group of all six-rowed invertible matrices. Since the symplectic matrices form a closed subgroup of this group, it follows that the Jacobian matrix of the transformation defined by (x_1, x_2, x_3) and (ξ_1, ξ_2, ξ_3) is symplectic, so that the transformation (2.5.17) is again canonical. This completes the proof. □

6 Regularisation of the solution of the three-body problem near a simple collision

We use the canonical transformation obtained in the previous section to uniformize the solution of three-body problem in the neigh bourhood of the singularity $t = t_1$ at which there is a simple collision. We recall that (x_1, x_2, x_3) and (x_4, x_5, x_6) are the relative coordinates of P_1 and P_2 with respect to P_3. We assume that the centre of gravity remains fixed at the
88 origin. (y_1, y_2, y_3) and (y_4, y_5, y_6) denote the components of momenta of P_1 and P_2 respectively. We have seen that $x_k, y_k,\ k = 1, \ldots, 6$, are obtained from the absolute coordinates q_k and the corresponding components of moments p_k by means of a canonical transformation. We now prove

Theorem 2.6.1. *The canonical transformation of the variables* (x_k, y_k) *to* (ξ_k, η_k)*,* $k = 1, 2, 3$*, defined by*

$$x_k = \eta^2 \xi_k - 2\eta_k \sum_{l=1}^{3} \xi_l \eta_l,\ y_k = \frac{\eta_k}{\eta^2},\ k = 1, 2, 3, \tag{2.6.1}$$

where $\eta^2 = \eta_1^2 + \eta_2^2 + \eta_3^2 \neq 0$*, can be extended to a canonical transformation of the twelve independent variables* (x_k, y_k) *to* (ξ_k, η_k)*,* $k = 1, \ldots, 6$*.*

Proof. By Theorem 2.5.1, (2.6.1) is a canonical transformation in the six variables $x_k, y_k, k = 1, 2, 3$. Let

$$A = (x_{k\xi_l}), B = (x_{k\eta_l}), C = (y_{k\xi_l}), D = (y_{k\eta_l}).$$

Then, $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, the Jacobian matrix of the transformation (2.6.1) is symplectic. We extend the transformation (2.6.1) to a transformation of (x_k, y_k) to $(\xi_k, \eta_k), k = 1, \ldots, 6$, by defining

$$x_k = \xi_k,\ y_k = \eta_k,\ k = 4, 5, 6. \tag{2.6.2}$$

Then the Jacobian matrix of the extended transformation is

$$M_1 = \begin{pmatrix} A & 0 & B & 0 \\ 0 & E_3 & 0 & 0 \\ 0 & 0 & D & 0 \\ 0 & 0 & 0 & E_3 \end{pmatrix}$$

where E_3 is the three-rowed unit matrix. Denoting by J the twelve-rowed square matrix $J = \begin{pmatrix} 0 & E_6 \\ -E_6 & 0 \end{pmatrix}$ where $E_6 = \begin{pmatrix} E_3 & 0 \\ 0 & E_3 \end{pmatrix}$, it is easy to **89**
verify that $M_1' J M_1 = J$, using the fact that M is symplectic. This proves the fact that (2.6.1) and (2.6.2) together define a canonical transformation which extends (2.6.1).

Now we recall that $s = \int\limits_{\tau}^{t} \dfrac{dt}{x}$ converges, as $t \to t_1$, to a finite limit $s_1 = \int\limits_{\tau}^{t_1} \dfrac{dt}{x}$. We consider the variables ξ_k, η_k as functions of the real variable s in the interval $0 \leq s < s_1$. We know that x_k, y_k are regular analytic functions of t in $\tau \leq t < t_1$ and hence, ξ_k, η_k, defined by (2.6.1) and (2.6.2), are regular analytic functions of s in $0 \leq s < s_1$. Then we have the following $\square$

Theorem 2.6.2. *The functions $\xi_k = \xi_k(s)$ and $\eta_k = \eta_k(s)$, $k = 1, \ldots, 6$, can be continued analytically as regular analytic functions of s to a neighbourhood of $s = s_1$.*

Proof. Since the canonical transformation defined by (2.6.1) and (2.6.2) is independent of the variable s, the Hamiltonian equations keep their form and therefore the Hamiltonian system $x_k' = F_{y_k}$, $y_k' = -F_{x_k}$, $k = 1, \ldots 6$, where

$$F = x(T - U - h) \tag{2.6.3}$$

goes over into the system

$$\xi_k' = \mathbb{F}_{\eta_k}, \eta_k' = -\mathbb{F}_{\xi_k},\ k = 1, \ldots, 6, \tag{2.6.4}$$

where $\mathbb{F}(\xi, \eta) = F(x, y) = x(T - U - h)$. In order to prove the the-
90 orem it is sufficient to prove that $\mathbb{F}\eta_k, \mathbb{F}_{\xi_k}$ are bounded regular analytic functions of the twelve variables (ξ, η) in some bounded closed region of 12-dimensional (ξ, η)-space. For this purpose we express $\mathbb{F}$ as a function of the variables (ξ_k, η_k). From the definition, we have

$$T = \frac{1}{2}\left(\frac{1}{m_1} + \frac{1}{m_3}\right)y^2 + \frac{1}{2}\left(\frac{1}{m_2} + \frac{1}{m_3}\right)\left(y_4^2 + y_5^2 + y_6^2\right)$$
$$+ \frac{1}{m_3}\left(y_1y_4 + y_2y_5 + y_3y_6\right),$$
$$U = \frac{m_1m_2}{r_{12}} + \frac{m_2m_3}{r_{23}} + \frac{m_1m_3}{x},$$

where $x^2 = x_1^2 + x_2^2 + x_3^2$, $y^2 = y_1^2 + y_2^2 + y_3^2$, $r_{23}^2 = x_4^2 + x_5^2 + x_6^2$ and $r_{12}^2 = (x_1 - x_4)^2 + (x_2 - x_5)^2 + (x_3 - x_6)^2$. On the other hand, we have from § 5 $y^2 = \dfrac{1}{\eta^2}$, $x = \xi\eta^2$, $xy^2 = \xi$ with $\xi^2 = \xi_1^2 + \xi_2^2 + \xi_3^2$, $\eta^2 = \eta_1^2 + \eta_2^2 + \eta_3^2$. Then we can write, by (2.6.1),

$$r_{12}^2 = \sum_{k=1}^{3}(x_k - \xi_{k+3})^2,\ r_{23}^2 = \sum_{k-1}^{3}\xi_{k+3}^2 \ \text{ where }\ x_k = \eta^2\xi_k - 2\eta_k\sum_{l=1}^{3}\xi_l\eta_l.$$

Denoting by b the positive constant $\dfrac{1}{2}(1/m_1 + 1/m_3)$, we have

$$\mathbb{F} = b\xi + \frac{1}{2}\left(\frac{1}{m_2} + \frac{1}{m_3}\right)\left(\eta_4^2 + \eta_5^2 + \eta_6^2\right)\xi\eta^2 + \frac{\xi}{m_3}\left(\eta_1\eta_4 + \eta_2\eta_5 + \eta_4\eta_6\right)$$
$$-h\xi\eta^2 - \left(\frac{m_1m_2}{r_{12}} + \frac{m_2m_3}{r_{23}}\right)\xi\eta^2 - m_1m_3. \tag{2.6.5}$$

In order to apply Cauchy's theorem on analytic continuation we need
91 to prove that $\mathbb{F}_{\xi_k}, \mathbb{F}_{\eta_k}$ are bounded regular functions of the twelve independent variables (ξ_k, η_k), $k = 1, \ldots, 6$. We have proved in § 5 that $xy^2 \to 2(m_1m_3)^2(m_1 + m_3)^{-1} = c > 0$. In other words,

$$\xi \to c > 0 \ \text{ as }\ s \to s_1. \tag{2.6.6}$$

Also $\eta_k = y_k y^{-2} \to 0$ as $t \to t_1$, i.e. as $s \to s_1$. Since $y \to \infty$ as $t \to t_1$, it follows that in a sufficiently small neighbourhood of $t = t_1$ we have $y \neq 0$. Let s_o be a number in $0 \leq s < s_1$ such that $y = y(s) \neq 0$ in $s_o \leq s < s_1$.

We know already that $\xi_k, \eta_k, k = 1, \ldots, 6$, are regular analytic functions of s in $0 \leq s < s_1$. By Theorem 2.3.2, we know that P_2 stays away from the colliding points P_1, P_3 and that its velocity components have finite limits as $t \to t_1$ and hence as $s \to s_1$. Then it follows that η_4, η_5 and η_6 have finite limits as $s \to s_1$. Moreover, all the absolute coordinates q_k tend to finite limits as $t \to t_1$, and hence $\xi_{k+3} = x_{k+3} = q_{k+3} - q_{k+6}$, $k = 1, 2, 3$, tend to finite limits as $s \to s_1$. By (2.6.6), $\xi \to c > 0$ as $s \to s_1$, and it will appear later that ξ_1, ξ_2, ξ_3 themselves tend to finite limits.

We choose the number s_o in the interval $0 \leq s < s_1$ so close to s_1 that for s in the interval $s_o \leq s < s_1, \xi = \xi(s)$ lies in the interval $c/2 \leq \xi(s) \leq 2c$. We have also seen that r_{12} and r_{23} remain bounded below by positive constants in $s_o \leq s < s_1$.

We now consider $\xi_k, \eta_k, k = 1, \ldots, 6$, as twelve independent vari- **92**
ables. Let D be a bounded closed region in the 12 dimensional (ξ, η)-space defined as follows. Let (ξ_1, ξ_2, ξ_3) be restricted to the annular region $D_1 : c/2 \leq \xi \leq 2c$ in 3 - dimensional ξ-space. Since ξ_4, ξ_5, ξ_6, η_4, η_5, η_6 have finite limits as $s \to s_1$, and $\eta_1, \eta_2, \eta_3 \to 0$ as $s \to s_1$, we can enclose the points $(\xi_4, \xi_5, \xi_6, \eta_1, \ldots, \eta_6)$, for s sufficiently close to s_1, in a bounded closed region D_2 in 9-dimensional space such that D_2 contains all the limit values of $\xi_4, \xi_5, \xi_6, \eta_1, \ldots, \eta_6$ in its interior. Moreover we can choose D_2 so small that r_{12}^{-1} and r_{23}^{-1} are bounded. Then we take $D = D_1 \times D_2$.

Consider the function $\mathbb{F}$ in the region D. $\mathbb{F}$ contains r_{12} and r_{23} in the denominator. On differentiation we observe that $\mathbb{F}_{\xi_k}$ is a function of (ξ_k, η_k) with ξ, r_{12}, r_{23} occuring in the denominator. Since ξ is bounded away from zero for $(\xi_1, \xi_2, \xi_3) \in D_1$ and r_{12}^{-1}, r_{23}^{-1} are bounded, it follows that the $\mathbb{F}_{\xi_k}$ are regular analytic functions of $\xi_k, \eta_k, k = 1, \ldots, 6$, in D. For the same reason the $\mathbb{F}_{\eta_k}$ are regular analytic in D. Consider the orbit $(\xi_k(s), \eta_k(s))$, $s_o \leq s < s_1 \cdot (\xi_k(s), \eta_k(s))$ is a curve in twelve-dimensional Euclidean space. If s_o is so chosen that the interval $[s_o, s_1]$ is sufficiently small, then this curve lies completely in the region D. Thus the hypotheses ensuring the existence of analytic continuation of solutions (Chapter 1, § 3) are satisfied for the system of differential equations

$$\xi'_k = \mathbb{F}_{\eta_k}, \eta'_k = -\mathbb{F}_{\zeta_k},\ k = 1, \dots 6.$$

93 It follows that the solution $\xi_k = \xi_k(s)$, $\eta_k = \eta_k(s)$ can be continued analytically on regular analytic functions of s in a neighbourhood of $s = s_1$. (Acutally it is possible to continue analytically even to a complex neighbourhood of $s = s_1$). This neighbourhood can be determined explicitly by obtaining estimates for the derivatives $\mathbb{F}_{\xi_k}$ and $\mathbb{F}_{\eta_k}$; it is quite straightforward to compute these estimates and we shall not do this. Thus it follows, in particular, that ξ_k tend to finite limits ξ_{kl} as $s \to s_1$ and the the η_k are regular analytic at $s = s_1$ with $\eta_k \to 0$ as $s \to s_1$, $k = 1, 2, 3$. This completes the proof of the theorem. □

Theorem 2.6.2 implies that we can expand ξ_k, η_k as power-series in $s - s_1$ in a neighbourhood of $s = s_1$. Substituting the expansions in the differential equations we obtain the coefficients of the power-series for $\xi_k(s)$ and $\eta_k(s)$. We consider first η_k, $k = 1, 2, 3$. Since $\xi_{\xi_k} = \dfrac{\xi k}{\xi}$, $k = 1, 2, 3$, we find on differentiating $\mathbb{F}$ with respect to ξ, that

$$\eta'_k = -\mathbb{F}_k = \frac{b\xi_k}{\xi} + \text{ terms vanishing for } s = s_1.$$

Since $\xi \to c$ and $\xi_k \to \xi_{k1}$ as $s \to s_1$, we can expand $\dfrac{\xi_k}{\xi}$ as a power series in $s - s_1$ in a neighbourhood of $s = s_1$ and obtain

$$\eta'_k = -\frac{b}{c}\xi_{k1} + \text{ terms of degree } \geq 1 \text{ in } s - s_1,\ k = 1, 2, 3.$$

94 Integrating from s to s_1, for s in a neighbourhood of s_1, and using the fact that $\eta_k \to 0$ as $s \to s_1$, i.e. $\eta_k(s_1) = 0$, we see that

$$\eta_k = -\frac{b}{c}\xi_{kl}(s - s_1) + \text{ terms of degree } \geq 2 \text{ in } s - s_1,\ k = 1, 2, 3. \quad (2.6.7)$$

But $\left(\frac{\xi(s_1)}{c}\right)^2 = \left(\frac{\xi_{11}}{c}\right)^2 + \left(\frac{\xi_{21}}{c}\right)^2 + \left(\frac{\xi_{31}}{c}\right)^2 = 1$. Squaring both sides of (2.6.7) and adding up for $k = 1, 2, 3$, we have

$$\eta^2 = b^2(s - s_1)^2 + \text{ terms of higher order, } b \neq 0 \tag{2.6.8}$$

which gives the following expansions for y_k, $k = 1, 2, 3$:

$$y_k = \frac{\eta_k}{\eta^2} = -\frac{1}{bc}\xi_{k1}(s - s_1)^{-1} + \text{ terms of degree } \geq 0 \text{ in } s - s_1. \tag{2.6.9}$$

This shows that at least one of the velocity components of P_1 and P_3 has a simple pole at $s = s_1$ and hence becomes inifinite of the order of $(s - s_1)^{-1}$ as $s \to s_1$. On the other hand, we have

$$\xi_k(x) = \xi_{k1} + \text{ terms of degree } \geq 1 \text{ in } s - s_1,\ k = 1, 2, 3, \tag{2.6.10}$$

so that

$$\xi(s) = c + \text{ terms of degree } \geq 1 \text{ in } s - s_1, \tag{2.6.11}$$

where $c^2 = \xi_{11}^2 + \xi_{21}^2 + \xi_{31}^2$. Substituting (2.6.7), (2.6.8) and (2.6.10) in (2.6.11) we obtain

$$\begin{aligned} x_k &= \eta^2\xi_k - 2\eta_k \sum_{l=1}^{3} \xi_l\eta_l \\ &= b^2\xi_{k1}(s - s_1)^2 + \ldots - 2(-\frac{b}{c}\xi_{k1}(s - s_1) + \ldots)(-bc(s - s_1) + \ldots) \end{aligned}$$

or **95**

$$x_k = -b^2\xi_{k1}(s - s_1)^2 + \text{ terms of degree } \geq 3 \text{ in } (s - s_1),\ k = 1, 2, 3. \tag{2.6.12}$$

Squaring and summing over $k = 1, 2, 3$, we get

$$x = b^2c(s - s_1)^2 + \text{ terms of degree } \geq 3 \text{ in } s - s_1. \tag{2.6.13}$$

By the definition of s we have $t' = \frac{dt}{ds} = x$. Integrating this over a sufficiently small interval (s, s_1) we get

$$t - t_1 = \frac{b^2c}{3}(s - s_1)^3 + \text{ terms of degree } > 3 \text{ in } (s - s_1). \tag{2.6.14}$$

Since $b \neq 0$, $c \neq 0$, we can invert this power-series in a neigh bourhood of $s = s_1$ and obtain

$$s_1 - s = \left(\frac{3}{b^2 c}(t_1 - t)\right)^{1/3} + \text{ terms of higher order in } (t_1 - t)^{\frac{1}{3}}. \quad (2.6.15)$$

As a consequence of (2.6.15) and the expansions (2.6.7) and (2.6.10) for ξ_k, η_k, it follows that ξ_k, η_k have power-series expansions in $(t - t_1)^{1/3}$ in some neighbourhood of $t = t_1$. Similarly the expansion (2.6.12) for x_k shows that x_k also has a power-series expansion in $(t - t_1)^{1/3}$ in a neighbourhood of $t = t_1$, $k = 1, 2, 3$. But from (2.6.9) we see that

$$\begin{aligned} y_k = \frac{\eta_k}{\eta^2} &= -\frac{1}{bc}\xi_{k1}(s - s_1)^{-1} + \dots \\ &= \frac{1}{bc}\xi_{k1}\left(\frac{3}{b^2 c}(t - t_1)\right)^{-\frac{1}{3}} + \dots, k = 1, 2, 3, \end{aligned}$$

96 contains also negative powers of $(t - t_1)^{1/3}$. This proves our conjecture that $t = t_1$ is an algebraic branch-point for some coordinate function $x_k, k = 1, 2, 3$, and that there are three sheets at the branch-point $t = t_1$.

As we mentioned earlier, Weierstrass had already asserted the existence of power-series expansions for the solution of the three-body problem in cube-roots of $t - t_1$. However, a proof was given explicitly for the first time by Sundman. Sundman's method was different from the one we have described. He did not use canonical transformations. They were first used by Levi-Civita. Sundman also showed that if we introduce the variable s, then we can get analytic continuations of the solutions even beyond $s = s_1$. As we pass the singularity $s = s_1$, it follows from (2.6.14), since the power-series starts with an odd power of $s - s_1$, that we can go beyond $t = t_1$ through real values. Because of (2.6.12), the collision thus means only a reflection of the colliding masses. The same system of differeential equations continues to be satisfied after the collision and so we are led back to the old problem. We can therefore continue the orbit and there are only two possibilities. Either no further collision occurs, or there is a next singularity t_2 at which there is a collision which cannot be a general collision if $\lambda^2 + \mu^2 + \nu^2 > 0$. This second collision may, however, not be between the same two masses as before.

And this process may continue. Suppose that $t_1, t_2, \ldots$ are the times of the successive simple collisions. Then, either this sequence is finite, in which case all the coordinates are regular analytic functions beyond the 97
last singularity, or there is an infinite sequence of simple collisions. In the latter case Sundman proved that $t_n \to \infty$ necessarily as $n \to \infty$. This is done in the following way. Suppose, if possible, that t_n has a finite limit point t_∞ as $n \to \infty$. We know that the potential function $U(t) \to \infty$ as $t \to t_n$ for each $n = 1, 2, \ldots$, and that $U(t)$ is finite between any two successive collision times $t_{n-1}, t_n, n = 2, 3, \ldots$. We now assert that $U(t) \to \infty$ as $t \to t_\infty$. For, if not, let $U(t) < A$ for t arbitrarily near t_∞. Then by the Cauchy existence theorem the coordinate functions $q(t)$ are regular analytic functions in a neighbourhood $|t - t_o| < B$ of t_o where B is a constant depending on A, the masses and the energy constant. Hence the $q(t)$ are regular at $t = t_\infty$ and so $U = T - h$ is regular at $t = t_\infty$. However, this is not possible since in any neighbourhood of t_∞ there exists a t_n at which U becomes infinite. Thus $U(t) \to \infty$ as $t \to t_\infty$. Consequently, $r(t) = \min(r_{12}, r_{23}, r_{13})$ tends to zero as $t \to t_\infty$. Now, by the Lagrange formula, $\ddot{\sigma} > 0$ in a sufficiently small interval $t_o \leq t < t_\infty$. However, $\ddot{\sigma}$ is infinite at each $t = t_n$, $n = 1, 2, \ldots$. We have seen that $\dot{\sigma}$ is continuous from the left at $t = t_1$ and we see similarly this is so at each $t = t_n$, and we conclude from our earlier discussion that it is also continuous on the right at each $t = t_n$. Thus $\dot{\sigma}$ is continuous and monotone increasing in $t_o \leq t < t_\infty$. From this it follows as before that σ has a positive lower bound in $t_o \leq t < t_\infty$ if $\lambda^2 + \mu^2 + \nu^2 > 0$. And now repeating our earlier argument we see that as $t \to t_\infty$, exactly one side of 98
the triangle P_1, P_2, P_3, say $r(t) = r_{13}$, tends to zero, and the other sides remain bounded away from zero. By the continuity of the distance functions we see that the collision is always between the same pair of points P_1, P_3 for all large n and hence we can take $r(t) = r_{13}$ for all large n. By our earlier arguments we can uniformize the coordinate functions $q(t)$ near $t = t_n$ for each $n = 1, 2, \ldots$ by means of the uniformising variable

$$s_n = \int_{t_o}^{t_n} \frac{dt}{r(t)},$$

the integral converging for each n. Repeating our earlier argument we see that $\dot{\sigma}$ is bounded as $t \to t_\infty$ and the integral

$$s_\infty = \int_{t_o}^{t_\infty} \frac{dt}{r(t)}$$

converges and so s_n tends to a finite limit s_∞ as $t \to t_\infty$. We can then prove that the coordinate functions $q(t)$ are regular analytic functions of s in a neighbourhood of $s = s_\infty$. Moreover, s and t are regular functions of each other and $\dfrac{dt}{ds} = r(t)$. Therefore $r(t) = r(t(s))$ is regular in a neighbourhood of $s = s_\infty$. But $r(t) = 0$ for each t_n, so that the analytic function $r(t)$ has an infinity of zeros in a neighbourhood of $s = s_\infty$. Thus $r(t) \equiv 0$ in a neighbourhood of $s = s_\infty$ and this is impossible. This
99 proves the assertion that the times of successive simple collisions t_n do not have a finite limit point.

Hereafter we shall deal with the general collision at $t = t_1$, in which case the singularity terms out, in general, to be an essential singularity.

Chapter 3

The three-body problem: general collision

1 Asymptotic estimates

In this chapter we shall be concerned with the problem of determining 100
the nature of the first singularity of the three-body problem when there is a general collision, that is when all the mass-points collide at $t = t_1$. We shall show that in this case, $t = t_1$ is in general an essential singularity for at least one of the coordinate functions of the masses.

Let us denote as before the coordinates of the three mass-points P_k by (x_k, y_k, z_k), $k = 1, 2, 3$, and their masses by m_k. Also let q denote any one of the nice coordinate functions $x_1, \ldots, z_3$. We recall that if U denotes the potential function

$$U = \frac{m_1 m_2}{r_{12}} + \frac{m_2 m_3}{r_{23}} + \frac{m_1 m_3}{r_{13}}, \tag{3.1.1}$$

then the equations of motion are given by

$$m\dot{q} = U_q. \tag{3.1.2}$$

We have the ten algebraic integrals associated with the system namely, the six integrals of the centre of gravity, the three integrals of angular momenta and the energy integral these are given by

$$\sum_{k=1}^{3} m_k x_k = \alpha t + \alpha', \sum_{k=1}^{3} m_k y_k = \beta t + \beta', \sum_{k=1}^{3} m_k z_k = \gamma t + \gamma';$$

$$\sum_{k=1}^{3} m_k (x_k \dot{y}_k - y_k \dot{x}_k) = \lambda, \sum_{k=1}^{3} m_k (y_k \dot{z}_k - z_k \dot{y}_k) = \mu,$$

101

$$\sum_{k=1}^{3} m_k (z_k \dot{x}_k - x_k \dot{z}_k) = \nu; \tag{3.1.3}$$

$$T - U = h. \tag{3.1.4}$$

We may assume, by changing the coordinates by linear functions of the variable t, that the centre of gravity remains fixed at the origin throughout the motion, so that

$$\sum_{k=1}^{3} m_k x_k = \sum_{k=1}^{3} m_k y_k = \sum_{k=1}^{3} m_k z_k = 0. \tag{3.1.5}$$

Let $t = t_1$ be the first singularity. So all the coordinates $q(t)$ are regular analytic functions in the interval $t_o \leq t < t_1$, and at least one coordinates ceases to be regular at t_1. Let ρ_k denote the distance of the point P_k from the centre of gravity 0:

$$\rho_k^2 = x_k^2 + y_k^2 + z_k^2, \ k = 1, 2, 3. \tag{3.1.6}$$

We introduced in Chapter 2 the moment of inertie

$$\sigma \equiv \sum_q m q^2 = \sum_{k=1}^{3} m_k \rho_k^2. \tag{3.1.7}$$

Differentiating this twice in succession with respect to we have

$$\frac{1}{2}\dot{\sigma} = \sum_q m q \dot{q}, \frac{1}{2}\ddot{\sigma} = \sum^{q} m \dot{q}^2 + \sum_q m q \ \ddot{q}. \tag{3.1.8}$$

102 Then we obtain the Lagrange formula

$$\frac{1}{2}\ddot{\sigma} = 2T - U, \tag{3.1.9}$$

which can be written, by means of (3.1.4), also as

$$\frac{1}{2}\ddot{\sigma} = T + h = U + 2h \tag{3.1.10}$$

By Theorem 2.2.1 we know that $U(t) \to \infty$ as $t \to t_1$, so that for t sufficiently close to t_1, $U+2h > 0$ and so $\ddot{\sigma} > 0$ and hence σ is a convex function of t and has a limit (non-negative) as $t \to t_1$. We may assume that σ is convex in the interval $\tau \leq t < t_1$ itself. We observe that $\sigma = 0$ if and only if all the coordinates q vanish, that is, $P_1, P_2, , P_3$ are all at 0. Since by assumption t_1 is the first singularity, there is no collision in the interval $\tau \leq t < t_1$ and so $\sigma(t) > 0$ in $\tau \leq t < t_1$. Further $\sigma(t)$ is a regular analytic function in this interval. We have seen in Chapter 2 that $\sigma(t)$ has a limit σ_1 as $t \to t_1 \cdot \sigma_1 > 0$ or ≥ 0 according as σ is monotone in monotone increasing or monotone decreasing in a small interval to the left of t_1, that is, according as $\dot{\sigma}(t) > 0$ or $\dot{\sigma}(t) < 0$ in this interval. If $\sigma_1 > 0$ there is only a simple collision at $t = t_1$. We have studied this cased in Chapter 2. So we shall consider only the case $\sigma_1 = 0$ (which can happen only when $\dot{\sigma}(t) < 0$) and study the nature of the singularity more closely in this case.

We shall now change our notation and introduce the variable $t_1 -$ 103
t instead of t. It is clear that the equations of motion (3.1.2) remain invariant under this change of variable. Then as t varies in the interval $\tau \leq t < t_1$, the variabel $t_1 - t$ varies in the interval $0 < t_1 - t \leq t_1 - \tau$ and $t_1 - t$ tends to 0 through decreasing values as $t \to t_1$ through increasing values. From now on we shall write t in place of the variable $t_1 - t$ and τ in place of $t_1 - \tau$. Thus in the new notation $0 < t \leq \tau$ and $t \to 0$ through decreasing real values. We consider the coordinates q as functions of the new variable t and write again $q = q(t)$. Now $U(t) \to \infty$ as $t \to 0$ and hence in a sufficiently small interval $0 < t \leq t_o$ with $0 < t_o \leq \tau$, we have $U(t) + 2h > 0$ so that $\ddot{\sigma}(t) > 0$ in $0 < t < t_o$. Therefore $\sigma(t)$ is again a convex function of t in $0 < t \leq t_o$. Moreover, $\ddot{\sigma}(t) > 0$ implies that $\dot{\sigma}(t)$ is monotone increasing as t decreases to 0 and is positive in $0 < t \leq t_o$.

We shall now study the asymototic behaviour of $\sigma(t)$ and $\dot{\sigma}(t)$ as $t \to 0$. First of all we have the following inequality:

Theorem 3.1.1. $\dot{\sigma}^2 \leq 8\sigma T$ *for* $0 < t \leq t_o$.

Proof. By the definition of σ and T we have $2\sigma T = \sum\limits_q mq^2 \sum_q m\dot{q}^2$. For arbitrary real numbers $\alpha_k, \beta_k, k = 1, \ldots, r$, we have the following identity of Lagrange (see Chapter 2, § 2):

$$\sum_{k=1}^{r} \alpha_k^2 \sum_{k=1}^{r} \beta_k^2 = \left(\sum_{k=1}^{r} \alpha_k \beta_k\right)^2 + \sum_{1 \leq k < l \leq r} (\alpha_k \beta_1 - \alpha_1 \beta_k)^2 .$$

Taking $r = 9$ and $\alpha_k = q_k \sqrt{m_k}, \beta_k = \dot{q}_k \sqrt{m_k}$, we obtain

$$2\sigma T = \left(\frac{1}{2}\dot{\sigma}\right)^2 + \sum_{1 \leq k < l \leq 9} m_k m_l (q_k \dot{q}_l - q_l \dot{q}_k)^2 \tag{3.1.11}$$

104 But

$$\sum_{1 \leq k < l \leq 9} m_k m_l (q_k \dot{q}_l - q_l \dot{q}_k)^2 \geq 0.$$

Hence from (3.1.11) we get $2\sigma T \geq (\frac{1}{2}\dot{\sigma})^2$, i.e. $\dot{\sigma}^2 \leq 8\sigma T$, which completes the proof. □

Theorem 3.1.2. *There is a positive constant κ such that*

$$\sigma(t) \sim \kappa t^{4/3} \ \ as \ \ t \to 0; \tag{3.1.12}$$

$$\dot{\sigma}(t) \sim \frac{4}{3}\kappa t^{1/3} \ \ as \ \ t \to 0. \tag{3.1.13}$$

Proof. By Theorem 3.1.1, $8\sigma T - \dot{\sigma}^2 \geq 0$. By the Lagrange formula (3.1.10), $\ddot{\sigma} = 2T + 2h$ and this can be written

$$\ddot{\sigma} = \frac{1}{4}(8\sigma T - \dot{\sigma}^2 + \dot{\sigma}^2)\sigma^{-1} + 2h,$$

and hence,

$$\ddot{\sigma} - \frac{1}{4}\dot{\sigma}^2\sigma^{-1} = \frac{1}{4}(8\sigma T - \dot{\sigma}^2)\sigma^{-1} + 2h. \tag{3.1.14}$$

Multiplying both sides by $\sigma^{-1/4}$ we find that

$$\left(\dot{\sigma}\sigma^{-1/4}\right) = \ddot{\sigma}\sigma^{-1/4} - \frac{1}{4}\dot{\sigma}^2\sigma^{-5/4} = \frac{1}{4}(8\sigma T - \dot{\sigma}^2)\sigma^{-5/4} + 2h\sigma^{-1/4}.$$

If we denote by $\sigma_o, \dot{\sigma}_o$ the values of $\sigma(t)$ and $\dot{\sigma}(t)$, at $t = t_o$, we obtain, on integration from t to t_o,

$$\dot{\sigma}_o\sigma_o^{-1/4} - \dot{\sigma}\sigma^{-1/4} = \frac{1}{4}\int_t^{t_o}(8\sigma T - \dot{\sigma}^2)\sigma^{-5/4}dt + 2h\int_t^{t_o}\sigma^{-1/4}dt. \quad (3.1.15)$$

Here $\sigma(t)$ and $\dot{\sigma}(t)$ are positive and hence $\dot{\sigma}\sigma^{-1/4} \geq 0$ as $t \to 0$, so it can only become $+\infty$ if it has no finite limit as $t \to 0$. Consequently the left side can at worst become $-\infty$ as $t \to 0$. On the right side, $(8\sigma T - \dot{\sigma}^2)\sigma^{-5/4} \geq 0$ by Theorem 3.1.1 and so $\int_t^{t_o}(8\sigma T - \dot{\sigma}^2)\sigma^{-5/4}dt \geq 0$. This has either a finite positive limit as $t \to 0$, or it tends to $+\infty$ as $t \to 0$. We shall, however, show that this integral converges and that $\dot{\sigma}\sigma^{-1/4}$ tends to a finite limit. For this it is sufficient to prove that the integral $\int_t^{t_o}\sigma^{-1/4}dt$ converges as $t \to 0$. **105**

We shall get a lower estimate for σ. In the following $\mu_1, \mu_2, \ldots$ denote positive constants which depend only on the masses m_1, m_2, m_3. Let $\mu = \min\limits_{k=1,2,3} m_k$, so that $\sigma \geq \mu(\rho_1^2 + \rho_2^2 + \rho_3^2)$ and by the Schwarz inequality this gives $\sigma \geq \frac{\mu}{9}(\rho_1 + \rho_2 + \rho_3)^2$. By the triangle inequality we have $r_{12} \leq \rho_1 + \rho_2$ and hence

$$\dot{\sigma} \geq \frac{\mu}{9}r_{12}^2, \quad \text{or} \quad \frac{1}{r_{12}} > \mu_1\sigma^{-1/2}. \quad (3.1.16)$$

This gives a lower estimate for the potential function: $U(t) > \mu_2\sigma^{-1/2}$. Moreover, $\frac{1}{2}\ddot{\sigma} = U + 2h$ so that, for t sufficiently near 0,

$$\ddot{\sigma}(t) > \mu_3\sigma^{-1/2}, \quad (3.1.17)$$

and on multiplication throughout by $2\dot{\sigma}$, $(\dot{\sigma}^2)^{\cdot} \leq 4\mu_3(\sigma^{1/2})^{\cdot}$. Integrating this from 0 to t we have

$$\dot{\sigma}^2 - \dot{\sigma}(0)^2 \geq \mu_4(\sigma^{1/2} - \sigma(0)^{1/2}).$$

106 Since $\sigma \to 0$ as $t \to 0$, $\sigma(0) = 0$ and further, $\dot{\sigma}(0)^2 \geq 0$, so that $\dot{\sigma}^2 \geq \mu_4\sigma^{1/2}$, and hence

$$\dot{\sigma} \geq \mu_5\sigma^{1/4}. \tag{3.1.18}$$

Then $(\sigma^3/4)^{\cdot} = \frac{3}{4}\dot{\sigma}\sigma^{-1/4} \geq \frac{3}{4}\mu_5$, which on integration from 0 to t, with $\sigma(0) = 0$, gives $\sigma^{3/4} \geq \mu_6 t$, or

$$\sigma(t) \geq \mu_7 t^{4/3},\ 0 < t \leq t_o. \tag{3.1.19}$$

Thus we obtain an upper estimate for the integral $\int\limits_t^{t_o} \sigma^{-1/4}dt$:

$$\int\limits_t^{t_o} \sigma^{-1/4}dt \leq \mu_8 \int\limits_t^{t_o} t^{-1/3}dt,$$

and the last integral converging as $t \to 0$, the convergence of $\int\limits_t^{t_o} \sigma^{-1/4}dt$ follows.

Now we shall show that $\sigma(t)$ actually behaves like $t^{4/3}$ asymptotically as $t \to 0$. Consider the indentity (3.1.15). Since the second integral on the right converges, it follows that $\dot{\sigma}\sigma^{-1/4}$ tends to a finite limit $a \geq 0$ as $t \to 0$. So $(\sigma^{3/4})^{\cdot} = \frac{3}{4}\dot{\sigma}\sigma^{-1/4} \to \frac{3}{4}a$ as $t \to 0$. In other words. $(\sigma^{3/4})^{\cdot} = \frac{3}{4}a + o(1)$ as $t \to 0$. Integrating this from 0 to t we see that

$$\sigma^{3/4} = \frac{3}{4}\text{ at } + o(t),$$

107 which implies that, as $t \to 0$,

$$\sigma(t) \sim (\frac{3}{4}a)^{4/3} t^{4/3}. \tag{3.1.20}$$

The constant a is necessarily positive as $a = 0$ would imply that $\sigma(t) \to 0$ more rapidly than $t^{4/3}$ as $t \to 0$, which is not possible because of (3.1.19). So $a > 0$; let $\kappa = (\frac{3}{4}a)^{4/3}$. Then we have $\sigma \sim \kappa t^{4/3}$ as $t \to 0$, which proves the first assertion. Further, from $\frac{3}{4}\dot{\sigma}\sigma^{-1/4} = \frac{3}{4}a + o(1)$, we have

$$\dot{\sigma} = a\sigma^{1/4} + o(\sigma^{1/4}) \sim \frac{4}{3}\kappa^{3/4}\kappa^{1/4}t^{1/3},$$

that is, $\dot{\sigma} \sim \frac{4}{3}\kappa t^{1/3}$ as $t \to 0$. This completes the proof of the theorem. □

We remark that the asymptotic estimate for $\dot{\sigma}$ proved directly in Theorem 3.1.2 is an improvement over the asymptotic estimate $\sigma \sim \kappa t^{4/3}$, in the following sense. If we could "differentiate" the asymptotic relation (3.1.12) for σ with respect to t, we would have obtained the asymptotic relation (3.1.13) for $\dot{\sigma}$. But in general such a differentiation is not permissible and so the direct proof above is an improvement.

We conjecture that we can again "differentiate" the relation (3.1.13) formally and obtain an asymptotic estimate for $\ddot{\sigma}$ in the form

$$\ddot{\sigma} = \frac{4}{9}\kappa t^{-2/3} \text{ as } t \to 0. \tag{3.1.21}$$

We shall see later that (3.1.21) in fact holds. For proving this we shall 108
again use the Lagrange formula $\frac{1}{2}\ddot{\sigma} = U + 2h$. So we proceed first to determine the asymptotic behaviour of $U(t)$ itself as $t \to 0$.

Consider the function

$$g(t) = (8\sigma T - \dot{\sigma}^2)t^{-2/3}. \tag{3.1.22}$$

In the course of the proof of Theorem 3.1.2 we have shown that the integral

$$\int_t^{t_o} (8\sigma T - \dot{\sigma}^2)\sigma^{-5/4} dt$$

converges as $t \to 0$. This means that the integral $\int\limits_t^{t_o} g(t)t^{2/3}\sigma^{-5/4}dt$ converges as $t \to 0$. But we have proved above that $\sigma \sim \kappa t^{4/5}$ and so $\sigma^{-5/4} \sim \kappa^{-5/4}t^{-5/3}$. Hence

$$\int\limits_t^{t_o} g(t)t^{2/3}\sigma^{-5/4}dt = \int\limits_t^{t_o} g(t)(\kappa^{-5/4}t^{-1} + o(t^{-1}))dt$$

converges as $t \to 0$. In particular, the integral

$$\int\limits_t^{t_o} g(t)\frac{dt}{t} \tag{3.1.23}$$

converges as $t \to 0$. But by Theorem 3.1.1, $g(t) \geq 0$ and therefore it follows that

$$\underline{\lim}_{t\to 0}\, g(t) = 0,$$

109 for $\underline{\lim}_{t\to 0}\, g(t) > 0$ would imply that the integral $\int\limits_t^{t_o} g(t)\frac{dt}{t}$ diverges. We shall next prove that

$$\overline{\lim}_{t\to 0} g(t) = 0.$$

Suppose if possible that $\overline{\lim}_{t\to 0} g(t) > 3\epsilon$, $0 < \epsilon < 1$. We shall prove that this leads to a contradiction. By the continuity of $g(t)$ in the interval $0 < t \leq \tau$, we can find a decreasing sequence of numbers $\tau \geq t_1 > t_2 > \ldots > 0$ such that

$$\epsilon \leq g(t) \leq 3\epsilon,\ t_{2k} \leq t \leq t_{2k-1}, \tag{3.1.24}$$

$$g(t_{2k}) = \epsilon,\ g(t_{2k-1}) = 3\epsilon. \tag{3.1.25}$$

By Theorem 3.1.2 we know that there exists a positive number κ depending only on the three masses such that $\sigma(t) \sim \kappa t^{4/3}$ and $\dot{\sigma}(t) \sim \frac{4}{3}\kappa t^{1/3}$ as $t \to 0$. Hence

$$\sigma(t) = \kappa t^{4/3}(1 + \delta_o(t)),\ \dot{\sigma}(t) = \frac{4}{3}\kappa t^{1/3}(1 + \delta_1(t)), \tag{3.1.26}$$

where $\delta_o(t), \delta_1(t) \to 0$ as $t \to 0$. If an asymptotic estimate for T were known we could then get an estimate for $8\sigma T t^{-2/3}$ and hence also for $g(t)$. But we do not have such an estimate for T as yet. However, we can get an upper estimate for T as follows. From the definition of g we have

$$T = \frac{1}{8}(g(t)t^{2/3} + \dot{\sigma}^2)\sigma^{-1}.$$

By (3.1.24), $g(t) \leq 3\epsilon$, $t_{2k} \leq t \leq t_{2k-1}$. On the other hand, by (3.1.26), **110**

$$\sigma^{-1} = \kappa^{-1}t^{-4/3}(1+\delta_o(t))^{-1},\ \dot{\sigma}^2 = (\frac{4}{3}\kappa)^2 t^{2/3}(1+\delta_1(t))^2.$$

Hence we obtain

$$T \leq \frac{1}{8}(3\epsilon t^{2/3} + (\frac{4}{3}\kappa)^2 t^{2/3}(1+\delta_1(t))^2)\kappa^{-1}t^{-4/3}(1+\delta_o(t))^{-1}.$$

It follows that $T \leq$ constant. $t^{-2/3}$, that is

$$T = 0(t^{-2/3}) t_{2k} \leq t \leq t_{2k-1}, k \to \infty. \tag{3.1.27}$$

Since $T = \dfrac{1}{2}\sum\limits_q m\dot{q}^2$, (3.1.27) implies, in particular, that $\dot{q}^2 = 0(t^{-2/3})$ as $t \to 0$, $t_{2k} \leq t \leq t_{2k-1}$, so that we have an estimate for the velocity components of P_1, P_2, P_3:

$$\dot{q} = 0(t^{-1/3}) \text{ as } t \to 0. \tag{3.1.28}$$

In view of the energy integral $T - U = h$ then we have

$$U = T - h = 0(t^{-2/3}) \text{ as } t \to 0. \tag{3.1.29}$$

By the definition of U, (3.1.29) implies that

$$r_{kl}^{-1} = 0(t^{-2/3}) \text{ as } t \to 0. \tag{3.1.30}$$

From these estimate we can get upper estimates for the derivatives $\dot{U}$ and $\dot{T}$ in the following way. We have

$$\dot{U} = -\sum_{1\leq k<l\leq 3} \frac{m_k m_l}{r_{kl}^3}(x_k - x_l)(\dot{x}_k - \dot{x}_l).$$

But $\dfrac{|x_k - x_l|}{r_{kl}} \leq 1$ and so we have 111

$$|\dot{U}| \leq \sum_{1\leq k<l\leq 3} \frac{m_k m_l}{r_{kl}^2}(|\dot{x}_k| + |\dot{x}_l|).$$

Then the estimates (3.1.28) and (3.1.30) show that

$$\dot{U} = 0(t^{-5/3}) \text{ as } t \to 0. \tag{3.1.31}$$

Once again, by the energy integral, $\dot{T} = \dot{U}$, so that

$$\dot{T} = 0(t^{-5/3}) \text{ as } t \to 0. \tag{3.1.32}$$

The asymptotic formula (3.1.26) together with the estimate (3.1.27) and (3.1.32) enables us to calculate the total variation of $g(t)$ in the interval $t_{2k} \leq t \leq t_{2k-1}$. In fact, on differentiation with respect to t we have

$$(T\sigma t^{-2/3})^{\cdot} = \dot{T}\sigma t^{-2/3} + T\dot{\sigma} t^{-2/3} - \frac{2}{3} T\sigma t^{-5/3} = 0(t^{-1}), \text{ as } t \to 0$$
$$\text{and} \quad t_{2k} \leq t \leq t_{2k-1}.$$

Hence we have the inequality

$$8(T\sigma t^{-2/3})^{\cdot} \leq \mu t^{-1} \text{ as } t \to 0,$$

and on integration from t_{2k} to t_{2k-1} we have

$$8 \int_{t_{2k}}^{t_{2k-1}} (T\sigma t^{-2/3})^{\cdot} dt \leq \mu \int_{t_{2k}}^{t_{2k-1}} \frac{dt}{t} \text{ as } k \to \infty. \tag{3.1.33}$$

112 On the other hand, using the asymptotic estimate (3.1.26) for $\dot{\sigma}(t)$, we find that the total variation of $\dot{\sigma}^2 t^{-2/3}$ in the interval $t_{2k} \leq t \leq t_{2k-1}$ is $o(1)$ as $k \to \infty$. Thus there exists a positive integer k_o such that for $k \geq k_o$, the variation of $\dot{\sigma}^2 t^{-2/3}$ in the interval $t_{2k} \leq t \leq t_{2k-1}$ is smaller than ϵ and also the estimate (3.1.33) for the variation of $T\sigma t^{-2/3}$ holds. So the variation of $g(t)$ in the interval $t_{2k} \leq t \leq t_{2k-1}$ is estimated by

$$2\epsilon = g(t_{2k-1}) - g(t_{2k}) \leq \epsilon + \mu \int_{t_{2k}}^{t_{2k-1}} \frac{dt}{t}, \ k \geq k_o.$$

Hence $\int\limits_{t_{2k}}^{t_{2k-1}} \frac{dt}{t} \geq \frac{\epsilon}{\mu}$, from which it follows that for $k \geq k_o$

$$\int\limits_{t_{2k}}^{t_{2k-1}} g(t)\frac{dt}{t} \geq \frac{\epsilon^2}{\mu}.$$

Since corresponding to $k \geq k_o$ there are infinitely many disjoint intervals $t_{2k} \leq t \leq t_{2k-1}$, each of which gives a contribution exceeding ϵ^2/μ to the integral, it follows that $\int\limits_0^\tau g(t)\frac{dt}{t}$ diverges, which is a contradiction. This means that necessarily $\overline{\lim\limits_{t\to 0}} g(t) = 0$. and hence $g(t) \to 0$ as $t \to 0$, which proves the required assertion.

As a consequence of (3.1.22) and the fact that $g(t) \to 0$ as $t \to 0$, we have

$$8\sigma T - \dot{\sigma}^2 = 0(t^{2/3}) \text{ as } t \to 0, \tag{3.1.34}$$

and hence 113

$$T \sim \frac{1}{8}(\frac{4}{3}\kappa t^{1/3})^2(\kappa t^{4/3})^{-1} = \frac{2}{9}\kappa t^{-2/3}, \text{ as } t \to 0. \tag{3.1.35}$$

By the energy integral $U = T - h$ it follows now that

$$U \sim \frac{2}{9}\kappa t^{-2/3} \text{ as } t \to 0. \tag{3.1.36}$$

We have already proved in Chapter 2 the theorem of Sundman that if there is a general collision at $t = 0$, then the constants of angular momenta λ, μ, ν all vanish. This can also be proved with the help of the estimates we have obtained, in the following way.

Theorem 3.1.3 (Sundman)**.** *If there is a general collision at $t = 0$, then $\lambda = \mu = \nu = 0$.*

Proof. We denote by $q_1, \ldots, q_9$ the nine coordinates $x_1, \ldots, z_3$, and by $\mu_1, \ldots, \mu_9$ the corresponding masses. Taking $\alpha_k = q_k\sqrt{\mu_k}$ and $\beta_k = \dot{q}_k\sqrt{\mu_k}$ in the Lagrange identity

$$\sum_{k=1}^{9}\alpha_k^2\sum_{k=1}^{9}\beta_k^2 = \left(\sum_{k=1}^{9}\alpha_k\beta_k\right)^2 + \sum_{1\leq k<l\leq q}(\alpha_k\beta_l - \alpha_l\beta_k)^2,$$

we obtain

$$8\sigma T - \dot{\sigma}^2 = 4 \sum_{1\le k<l\le 9} \mu_k \mu_1 (q_k \dot{q}_l - q_l \dot{q}_k)^2.$$

Since by (3.1.34), $8\sigma T - \dot{\sigma}^2 = o(t^{2/3})$ as $t \to 0$, it follows that

$$\sum_{1\le k<l\le 9} \mu_k \mu_l (q_k \dot{q}_l - q_l \dot{q}_k)^2 = o(t^{2/3}) \text{ as } t \to 0.$$

114 Since all the quantities $\mu_k, q_k, \dot{q}_k$ are real, we see immediately that

$$q_k \dot{q}_l - q_l \dot{q}_k = o(t^{1/3}) \text{ as } t \to 0,\ k,\ l = 1, \ldots, 9,\ k \neq l.$$

(This estimate would naturally not be valid in the complex case). If we change our notation and denote by p, q two distinct coordinates $x_1, \ldots, z_3$, then we can write

$$p\dot{q} - q\dot{p} = o(t^{1/3}) \text{ as } t \to 0. \tag{3.1.37}$$

We recall that the integrals of angular momenta are given by

$$\sum_{k=1}^{3} m_k (x_k \dot{y}_k - y_k \dot{x}_k) = \lambda, \sum_{k=1}^{3} m_k (y_k \dot{z}_k - z_k \dot{y}_k) = \mu,$$
$$\sum_{k=1}^{3} m_k (z_k \dot{x}_k - x_k \dot{z}_k) = \nu.$$

If there is a general collision at $t = 0$, then $\sigma(t) \to 0$ as $t \to 0$ and the estimates (3.1.27) hold. Consequently, taking for p, q the coordinates $x_k, y_k; y_k, z_k; z_k, x_k$ in turn, we see that $\lambda = \mu = \nu = 0$, and this completes the proof. □

We remark that the converse of Theorem 3.1.3 is not in general true, that is to say, $\lambda = \mu = \nu = 0$ does not necessarily imply that there is a general collision at the singularity $t = 0$.

2 The limiting configuration at a general collision

We shall now prove a result due to Sundman that in the case of a general collision the three mass-points always lie in a fixed plane through the
115 centre of gravity (which is assumed to be fixed fixed at the origin). For this purpose we shall proceed as follows.

It may happen that at the initial time $t = \tau$ the three point P_1, P_2, P_3 lie on the same straight line. We assume for the moment that this is not the case. Hence at time $t = \tau$ the area of the triangle $P_1P_2P_3$ is different from zero. This plane can be taken to be the (x, y)-plane, by means of an orthogonal transformation, if necessary, applied to the plane determined by P_1, P_2, P_3. We verify first that the differential equations of motion remain invariant under a fixed orthogonal transformation. Let $A = (a_{kl})$ denote the three-rowed matrix of the orthogonal transformation. Then we have

$$\sum_{j=1}^{3} a_{kj}a_{lj} = \delta_{kl},\ k, l = 1, 2, 3, \tag{3.2.1}$$

where $\delta_{kk} = 1$ and $\delta_{kl} = 0$ if $k \neq l$. Let (x_k, y_k, z_k) denote the original coordinates of the point P_k and (X_k, Y_k, Z_k) its coordinates after the orthogonal transformation. Then,

$$\begin{aligned} X_k &= a_{11}x_k + a_{12}y_k + a_{13}z_k, \\ Y_k &= a_{21}x_k + a_{22}y_k + a_{23}z_k, Z_k = a_{31}x_k + a_{32}y_k + a_{33}z_k. \end{aligned}$$

Differentiating X_k twice with respect to t and using the equations of motion $m_k\ddot{q}_k = U_{q_k}, q_k = x_k, y_k, z_k$, $k = 1, 2, 3$, we have

$$\begin{aligned} m_k\ddot{X}_k &= a_{11}m_k\ddot{x}_k + a_{12}m_k\ddot{y}_k + a_{13}m_k\ddot{z}_k \\ &= a_{11}U_{x_k} + a_{12}U_{y_k} + a_{13}U_{z_k}. \end{aligned} \tag{3.2.2}$$

On the other hand, we have, by the chain-rule,

$$U_{x_k} = U_{X_k}(X_k)_{x_k} + U_{Y_k}(Y_k)_{x_k} + U_{Z_k}(Z_k)_{x_k} = a_{11}U_{X_k} + a_{21}U_{Y_k} + a_{31}U_{Z_k},$$

and we have similar relations for U_{y_k} and U_{z_k}. Therefore, using (3.2.1) 116
and (3.2.2) we get $m_k\ddot{X}_k = U_{X_k}$, $k = 1, 2, 3$, and similarly, $m_k\ddot{Y}_k =$

U_{Y_k}, $m_k\ddot{Z}_k = U_{Z_k}$, $k = 1, 2, 3$. Thus the differential equations remain unchanged by the orthogonal transformation. Moreover, if the integrals of angular momenta λ, μ, ν, vanish in the original coordinate system, they vanish also in the new coordinate system. This follows by direct computation. For instance,

$$\begin{aligned}\sum_{k=1}^{3} m_k(X_k\dot{Y}_k - Y_k\dot{X}_k) = (a_{11}a_{22} - a_{12}a_{21}) \sum_{k=1}^{3} m_k(x_k\dot{y}_k - y_k\dot{z}_k)+ \\ + (a_{12}a_{23} - a_{13}a_{22}) \sum_{k=1}^{3} m_k(y_k\dot{z}_k - z_k\dot{y}_k)+ \\ + (a_{13}a_{21} - a_{11}a_{23}) \sum_{k=1}^{3} m_k(z_k\dot{x}_k - x_k\dot{z}_k) = 0.\end{aligned}$$

Then we have the following

Theorem 3.2.1 (Sundman)**.** *If the centre of gravity remains fixed at the origin and there is a general collision at $t = 0$, then the three mass-points P_1, P_2, P_3 remain in a fixed plane throughout the motion.*

Proof. Suppose that P_1, P_2, P_3 are not in the same straight line at $t = \tau$. We perform an orthogonal transformation and assume that P_1, P_2, P_3 lie in the (x, y)-plane at $t = \tau$. Then $z_k(\tau) = 0$, $k = 1, 2, 3$. Since the area of the triangle formed by P_1, P_2, P_3 at $t = \tau$ is not zero, we have

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix} \neq 0 \text{ at } t = \tau. \tag{3.2.3}$$

117 Since the centre of gravity remains fixed at the origin, we have

$$\sum_{k=1}^{3} m_k\dot{z}_k = 0 \text{ at } t = \tau. \tag{3.2.4}$$

Moreover, since $z_k(\tau) = 0$, $k = 1, 2, 3$, and there is a general collision at $t = 0$, we have from the integrals of angular momentum ($\lambda = \mu = \nu = 0$),

$$\sum_{k=1}^{3} m_k x_k\dot{z}_k = 0, \ \sum_{k=1}^{3} m_k y_k\dot{z}_k = 0, \text{ at } t = \tau. \tag{3.2.5}$$

Equations (3.2.4) and (3.2.5) form a system of three linear equations satisfied by $\dot{z}_k, k = 1, 2, 3$, at $t = \tau$. Since the matrix of this system of linear equations has, by (3.2.3), a determinant $\neq 0$, it follows that $\dot{z}_k(\tau) = 0$, $k = 1, 2, 3$. From the equations of motion we have

$$m_1\ddot{z}_1 = U_{z_1} = \frac{m_1 m_2}{r_{12}^3}(z_2 - z_1) + \frac{m_1 m_3}{r_{13}^3}(z_3 - z_1),$$

$$\text{or} \qquad \ddot{z}_1 = \frac{m_2}{r_{12}^3}(z_2 - z_1) + \frac{m_3}{r_{13}^3}(z_3 - z_1),$$

and similarly,

$$\ddot{z}_2 = \frac{m_1}{r_{12}^3}(z_1 - z_2) + \frac{m_3}{r_{23}^3}(z_3 - z_2), \ddot{z}_3 = \frac{m_1}{r_{13}^3}(z_1 - z_3) + \frac{m_2}{r_{23}^3}(z_2 - z_3).$$

At the initial time $t = \tau$, $r_{12}(\tau)$, $r_{23}(\tau)$, $r_{13}(\tau) \neq 0$. But since $z_k(\tau) = 0$, **118**
it follows from the equations above that $\ddot{z}_k(\tau) = 0$, $k = 1, 2, 3$. Differentiating the equations successively and using the fact that $z_k, \dot{z}_k, \ddot{z}_k$ vanish at $t = \tau$, we find that all the derivatives of z_k vanish at $t = \tau$, $k = 1, 2, 3$. Since there is no collision in the interval $0 < t \leq \tau$, we know that all the coordinate functions $z_k(t)$ are regular analytic functions in $0 < t \leq \tau$. It then follows that $z_k(t) \equiv 0$ for $0 < t < \tau$, $k = 1, 2, 3$. We could also prove this fact directly without making use of the analyticity of z_k in $0 < t \leq \tau$. In fact, consider the system of differential equation $m_k\ddot{q}_k = U_{q_k}$, $q_k = x_k, y_k, z_k$, $k = 1, 2, 3$. We prove as before that $\dot{z}_k(\tau) = 0$. Then we use the fact that if we fix $x_k(\tau)$, $y_k(\tau)$, $z_k(\tau)$ and $\dot{x}_k(\tau), \dot{y}_k(\tau), \dot{z}_k(\tau)$, then this system of differential equations has a unique solution. If we now fix $x_k(t)$ and $y_k(t)$, the differential equations for z_k with initial conditions $z_k(\tau) = 0 = \dot{z}_k(\tau)$ is identically satisfied by $z_k(t) = 0$, $0 < t \leq \tau$, because the differential equations contain the differences $z_k - z_l$ in the numerator. Then by uniqueness the two solutions coincide.

Next we consider the case in which P_1, P_2, P_3 lie on a straight line at the initial time $t = \tau$. Choose this line as the x-axis and choose as (x, y)-plane the plane determined by this line and the direction of the velocity vector of P_3 at the initial time, that is, $\dot{z}_3 = 0$ at $t = \tau$. Thus we have $y_1(\tau) = y_2(\tau) = y_3(\tau) = 0$. Since $\lambda = \mu = \nu = 0$, the condition

$\dot{z}_3(\tau) = 0$ implies that $m_1\dot{z}_1(\tau)$ and $m_2\dot{z}_2(\tau)$ satisfy the homogeneous
119 linear equations $m_1\dot{z}_1(\tau)+m_2\dot{z}_2(\tau) = 0$ and $m_1x_1(\tau)\dot{z}_1(\tau)+m_2x_2(\tau)\dot{z}_2(\tau) = 0$. The matrix of this system of homogeneous linear equations is

$$\begin{pmatrix} x_1(\tau) & x_2(\tau) \\ 1 & 1 \end{pmatrix}.$$

Since there is no collision at $t = \tau$, it follows that $x_1(\tau) \neq x_2(\tau)$ and hence this matrix has a non-vanishing determinant. One then obtains $\dot{z}_k(\tau) = 0$, $k = 1, 2, 3$. Repeating the argument used in the earlier case one has $z_1 = z_2 = z_3 = 0$ for all t in $0 < t \leq \tau$. This completes the proof of the theorem. □

In view of Theorem 3.2.1 we may assume that P_1, P_2, P_3 remain in the fixed plane $z = 0$ throughout the motion, so that $z_k(t) = 0$, $k = 1, 2, 3$, for all $t, 0 < t \leq \tau$. We wish to determine the behaviour of the six coordinates $x_k, y_k, k = 1, 2, 3$, near $t = 0$. Let q denote any of these six coordinates. By Theorem 3.1.2 we have $\sigma = \sum_q mq^2 \sim \kappa t^{4/3}$ as $t \to 0$, which implies that $q = 0(t^{2/3})$ as $t \to 0$. One would conjecture that every one of the six coordinates q can be expanded as a power-series in the variable $t^{1/3}$, starting with the term $t^{2/3}$, in a neighbourhood of $t = 0$. This was the case when there was a simple collision, as show in Chapter 2. It is no longer so in the case of a general collision. However, one can get an expansion for q in the variable $t^{1/3}$, this time with irrational exponents.

If q denotes any of $x_k, y_k, k = 1, 2, 3$, we set

$$q = q^* t^{2/3}. \tag{3.2.6}$$

120 Since $q = 0(t^{2/3})$, we have $q^* = 0(1)$ as $t \to 0$. Differentiating (3.2.6) with respect to t one obtains

$$\dot{q} = \dot{q}^* t^{2/3} + \frac{2}{3} q^* t^{-1/3}. \tag{3.2.7}$$

Similarly, if p denotes a coordinate distinct from q, let $p = p^* t^{2/3}$, then $p^* = 0(1)$ and

$$\dot{p} = \dot{p}^* t^{2/3} + \frac{2}{3} p^* t^{-1/3}. \tag{3.2.8}$$

From (3.2.7) and (3.2.8) we obtain

$$p\dot{q} - q\dot{p} = (p^*\dot{q}^* - q^*\dot{p}^*)t^{4/3}.$$

By (3.1.31), $p\dot{q} - q\dot{p} = 0(t^{1/3})$ as $t \to 0$ and hence we have

$$p^*\dot{q}^* - q^*\dot{p}^* = o(t^{-1}) \text{ as } t \to 0. \tag{3.2.9}$$

We introduce the following notation. If f is a homogeneous function of degree m in the variables $q_1, \ldots, q_6$, let f^* denote the function of the variables $q_1^*, \ldots, q_6^*$ defined by the relation $f^*(q^*) = f(q^*)$. Then f and f^* are related by the equation $f = f^* t^{2m/3}$. Since σ is a homogeneous function of degree 2 in $q_1, \ldots, q_6$, we have $\sigma = \sigma^* t^{4/3}$. On the other hand, $\sigma \sim \kappa t^{4/3}$ and so we have $\sigma^* \sim \kappa$ as $t \to 0$, that is

$$\sigma^*(t) = \kappa + o(1) \text{ as } t \to 0. \tag{3.2.10}$$

From the relation $\sigma = \sigma^* t^{4/3}$ we have, by differentiation, 121

$$\dot{\sigma}(t) = \dot{\sigma}^* t^{4/3} + \frac{4}{3}\sigma^* t^{1/3}, \tag{3.2.11}$$

Again by Theorem 3.1.2, $\dot{\sigma} \sim \frac{4}{3}\kappa t^{1/3}$ and (3.2.10) and (3.2.11) imply that $\dot{\sigma}^* t^{4/3} = o(t^{1/3})$ as $t \to 0$, or

$$\dot{\sigma}^*(t) = o(t^{-1}) \text{ as } t \to 0. \tag{3.2.12}$$

Since q is regular analytic in $0 < t \leq \tau$, q^* is also regular analytic in this interval and on differentiating σ^*, as we may, we have

$$\frac{1}{2}\dot{\sigma}^* = \sum_q mq^*\dot{q}^*. \tag{3.2.13}$$

The estimate $q^* = 0(1)$ together with (3.2.9) and (3.2.13) implies

$$\begin{aligned}
\frac{1}{2}\dot{\sigma}^* p^* - \dot{p}^*\sigma^* &= \sum_q m(q^*\dot{q}^* p^* - \dot{p}^* q^{*^2}) \\
&= \sum_q mq^*(\dot{q}^* p^* - \dot{p}^* q^*) = o(t^{-1}) \text{ as } t \to 0.
\end{aligned}$$

Using once again $p^* = 0(1)$, $\dot{\sigma}^* = 0(t^{-1})$, the last formula gives $\dot{p}^*\sigma^* = o(t^{-1})$ as $t \to 0$, from which it follows, by (3.2.10), that

$$\dot{p}^* = o(t^{-1}) \text{ as } t \to 0.$$

We have seen that $q^* = qt^{-2/3} = 0(1)$ as $t \to 0$ and we want to determine the exact behaviour of $q^*(t)$ as $t \to 0$. For this purpose we
122 consider the triangle determined by the points (x_k^*, y_k^*), $k = 1, 2, 3$, which will be referred to hereafter as the "big triangle". We observe that the centre of gravity of the system with respect to the $*$-coordinates also remains fixed at the origin: in fact,

$$\sum_{k=1}^{3} m_k x_k^* = t^{-2/3} \sum_{k=1}^{3} m_k x_k = 0, \ \sum_{k=1}^{3} m_k y_k^* = t^{-2/3} \sum_{k=1}^{3} m_k y_k = 0.$$

All the coordinates q^* are bounded as $t \to 0$ and we expect that q^* will have finite limit values as $t \to 0$, so that the big triangle has a limiting position as $t \to 0$. This will be proved only at the end. At present we have the following

Theorem 3.2.2. *Let the centre of gravity remain fixed at the origin and let there be a general collision at $t = 0$. Then the figure of the big triangle has a limiting configuration as $t \to 0$, and this limiting configuration is either an equilateral triangle or a set of three collinear points.*

Proof. We shall, first of all, write down the equations of motion $m\ddot{q} = U_q$, $q = x_k, y_k$, $k = 1, 2, 3$, in terms of the variables q^*. By definition we have $U = \sum_{1 \le k < l \le 3} m_k m_l r_{kl}^{-1}$, and r_{kl} being a homogeneous function of degree 1 in q_k, U is a homogeneous function of degree -1 in the six variables $x_k, y_k, k = 1, 2, 3$. Then U_q is a homogeneous function of degree -2. Using the notation introduced earliear, $U_q = U^*_{q^*}(q^*)t^{-4/3}$. Differentiating $q = q^* t^{2/3}$ with respect to t, we obtain

$$\begin{aligned} \ddot{q} &= \ddot{q}^* t^{2/3} + \frac{4}{3}\dot{q}^* t^{-1/3} - \frac{2}{9} q^* t^{-4/3} \\ &= (\dot{q} t^{4/3})^{\cdot} t^{-2/3} - \frac{2}{9} q^* t^{-4/3} \end{aligned} \tag{3.2.14}$$

123 The equations of motion now become

$$(\dot{q}^* t^{4/3})^{\cdot} t^{-2/3} - \frac{2}{9} q^* t^{-4/3} = \ddot{q} = \frac{1}{m} U_q = \frac{1}{m} U^*_{q^*} t^{-4/3},$$

or

$$-\frac{2}{9} q^* + (\dot{q}^* t^{4/3})^{\cdot} t^{2/3} = \frac{1}{m} U^*_{q^*}. \tag{3.2.15}$$

We shall replace each term in (3.2.15) by its average over the interval $(t, 2t)$, $0 < t < 2t \leq \tau$. We shall first prove that

$$\frac{1}{t} \int\limits_t^{2t} (\dot{q}^* t^{4/3})^{\cdot} t^{2/3} dt = o(1) \text{ as } t \to 0. \tag{3.2.16}$$

In fact, integrating the left side by parts, we obtain

$$\begin{aligned} \frac{1}{t} \int\limits_t^{2t} (\dot{q}^* t^{4/3})^{\cdot} t^{2/3} dt &= \frac{1}{t} \left(\left[\dot{q}^* t^{4/3} \cdot t^{2/3} \right]_t^{2t} - \frac{2}{3} \int\limits_t^{2t} \dot{q}^* t^{4/3} \cdot t^{-1/3} dt \right) \\ &= \frac{1}{t} (o(t^{-1}) t^2 - o(t)), \end{aligned}$$

as $t \to 0$, since $\dot{q}^* = o(t^{-1})$ as $t \to 0$, and so the right side is 0(1) as $t \to 0$, which proves (3.2.16). Next, we consider the two terms in (3.2.15). If t_1 and t_2 are real numbers such that $t \leq t_1 < t_2 \leq 2t$, then we have

$$q^*(t_2) - q^*(t_1) = \int\limits_{t_1}^{t_2} (q^*)^{\cdot} dt.$$

Since $\dot{q}^* = o(t^{-1})$ as $t \to 0$ and $0 < t_2 - t_1 \leq t$, it follows that the right 124
side is 0(1) and therefore we see that

$$q^*(t_2) = q^*(t_1) + o(t) \text{ as } t \to 0 \tag{3.2.17}$$

We also have

$$U^*_{q_*}(t_2) - U^*_{q^*}(t_1) = \int\limits_{t_1}^{t_2} (U^*_{q^*})^{\cdot} dt.$$

But $(U^*_{q^*})^{\cdot} = \sum\limits_{p^*} U^*_{q^* p^*} \dot{p}^*$, where p^* denotes any one of x^*_k, y^*_k, $k = 1, 2, 3$. Once again, since U is homogeneous of degree -1 in q, $U = U^* t^{-2/3}$. By (3.1.36), $U \sim \dfrac{2}{9}\kappa t^{-2/3}$ as $t \to 0$ and so we have

$$U = \frac{2}{9}\kappa t^{-2/3}(1 + \delta(t)), \delta(t) \to 0 \text{ as } t \to 0.$$

From this it follows that $U^* = \dfrac{2}{9}\kappa(1 + \delta(t))$, that is,

$$U^* = 0(1) \text{ as } t \to 0. \tag{3.2.18}$$

If we denote by $r^*_{kl}(t)$ the sides of the big triangle, then

$$r^*_{kl}(t)^2 = (x^*_k - x^*_l)^2 + (y^*_k - y^*_l)^2,$$

and we deduce from (3.2.18) and the definition of U^* the estimate

$$(r^*_{kl})^{-1} = 0(1) \text{ as } t \to 0. \tag{3.2.19}$$

Since $U^* = \sum\limits_{1 \le k < l \le 3} m_k m_l r^{*-1}_{kl}$, on differentiation with respect to p^* and q^* in succession, and then using (3.2.19), we get

$$U^*_{p^* q^*} = 0(1) \text{ as } t \to 0.$$

125 Since $(p^*)^{\cdot} = o(t^{-1})$ as $t \to 0$, we conclude that $(U^*_{q^*})^{\cdot} = o(t^{-1})$, so that we have as before

$$U^*_{q^*}(t_2) = U^*_{q^*}(t_1) = \int\limits_{t_1}^{t_2} (U^*_{q^*})^{\cdot} dt.$$

This implies that

$$U^*_{q^*}(t_2) = U^*_{q^*}(t_1) + o(1) \text{ as } t \to 0. \tag{3.2.20}$$

We take the average of (3.2.15) over the interval $(t, 2t)$, and using (3.2.16), (3.2.17) and (3.2.20) we obtain, for every fixed t_1 in $t \le t_1 \le 2t$,

$$-\frac{2}{9} q^*(t_1) + o(1) = \frac{1}{m} U^*_{q^*}(t_1) + o(1) \;\text{ as }\; t \to 0.$$

Hence, for t sufficiently near 0 we have

$$-\frac{2}{9}q^*(t) = \frac{1}{m}U^*_{q^*}(t) + o(1) \text{ as } t \to 0. \tag{3.2.21}$$

This is no longer a system of differential equations, but a system of six algebraic equations satisfied asymptotically by the six coordinates $q^*(t) = x^*_k(t), y^*_k(t), k = 1, 2, 3$. The system (3.2.21) can now be used to determine the behaviour of q^* more closely as $t \to 0$.

We observe that the system (3.2.21) is left invariant by an orthogonal transformation of the variables $x^*_y, y^*_k, k = 1, 2, 3$. As in the case of the proof of Theorem 3.2.1 this can be verified by a direct computation of $U^*_{q^*}$, in terms of the new variables. An orthogonal transformation corresponds to a rotaion of the axes (in the plane of the $*$-coodinates). We apply an orthogonal transformation (depending on t) in the plane of **126**
motion and assume that the new x^*-axis is parallel to the direction of the vector P_3P_1 (at time t). Let $X_k = X_k(t)$, $Y_k = Y_k(t)$ be the new coordinates of the points $P_k, k = 1, 2, 3$, at time t. Then $Y_1 = Y_3$ by assumption. Writing down the equations (3.2.21) for the coordinates Y_1, Y_2, Y_3, we have

$$\begin{aligned}
-\frac{2}{9}Y_1 &= \frac{m_2}{R^3_{12}}(Y_2 - Y_1) + -\frac{m_3}{R^3_{13}}(Y_3 - Y_1) + o(1),\\
-\frac{2}{9}Y_2 &= \frac{m_1}{R^3_{12}}(Y_1 - Y_2) + -\frac{m_3}{R^3_{23}}(Y_3 - Y_2) + o(1),\\
-\frac{2}{9}Y_3 &= \frac{m_1}{R^3_{13}}(Y_1 - Y_3) + -\frac{m_2}{R^3_{23}}(Y_2 - Y_3) + o(1),
\end{aligned}$$

as $t \to 0$, where R_{kl} denotes $r^*_{kl}(t)$, which is clearly left invariant by the orthogonal transformation. One can also write down similar algebraic equations for X_1, X_2, X_3. Since $Y_1 = Y_3$ the preceding equations become

$$\begin{gathered}
-\frac{2}{9}Y_1 = \frac{m_2(Y_2 - Y_1)}{R^3_{12}} + o(1), \quad -\frac{2}{9}Y_2 = \left(\frac{m_1}{R^3_{12}} + \frac{m_3}{R^3_{23}}\right)(Y_1 - Y_2) + o(1),\\
-\frac{2}{9}Y_3 = \frac{m_2(Y_2 - Y_1)}{R^3_{23}} + o(1),
\end{gathered} \tag{3.2.22}$$

as $t \to 0$. It is not immediate that Y_1, Y_2, Y_3 have limit values as $t \to 0$. Since $|Y_k - Y_l| R_{kl}^{-1} \leq 1$ and $R_{kl}^{-2} = (r_{kl}^*)^{-2} = 0(1)$ as $t \to 0$ by (3.2.19), it follows from (3.2.22) that $Y_1(t)$, $Y_2(t)$, $Y_3(t)$ are all $0(1)$ as $t \to 0$. Hence, by the Weierstrass theorem, we can find a sequence of values of
127 t which tends to zero such that the corresponding sequence of values of $Y_k(t)$ converges to a finite limit value as $t \to 0$ through this sequence. We shall see later on that these limit values of $Y_k(t)$ are independent of the sequence of values of t chosen. We denote also the limit values of the coordinates $X_1(t), \ldots, Y_3(t)$ by $X_1, \ldots, Y_3$ respectively, the limit values of $R_{kl}(t)$ by R_{kl}, $k \neq l$, and we can then omit the error term $o(1)$ in (3.2.22). Once again, since $Y_1 = Y_3$, we get from (3.2.22),

$$(Y_2 - Y_1)\,(R_{12}^{-3} - R_{23}^{-3}) = 0,$$

which means that either $Y_2 - Y_1 = 0$ or $R_{12}^{-3} - R_{23}^{-3} = 0$.

Suppose for the moment that $R_{12}^{-3} - R_{23}^{-3} \neq 0$. Then $Y_1 = Y_2$ and hence $Y_1 = Y_2 = Y_3$. Since the centre of gravity remains fixed at the origin, $m_1 Y_1 + m_2 Y_2 + m_3 Y_3 = 0$ and so it follows that $Y_1 = Y_2 = Y_3 = 0$, which means that the three points represented by $(X_k(t), Y_k(t))$, $k = 1, 2, 3$, tend to points situated on a straight line, as $t \to 0$.

Suppose on the other hand that $Y_1 \neq Y_2$; then necesarily $R_{12} = R_{23}$. If the three points are not collinear in the limiting position, then one can interchange P_1, P_2, P_3 (which means on orthogonal transformation with matrix independent of t) and repeat this argument and get $R_{23} = R_{13}$.

Hence, only two possibilities can occur, namely, either the three points represented by (X_k, Y_k), $k = 1, 2, 3$, are collinear, or they lie at
128 the vertices of an equilateral triangle, as $t \to 0$. We shall refer to these alternatives as the *collinear case* and the *equilateral case* respectively. This is equivalent to saying that either all the angles at the vertices tend to $\dfrac{\pi}{3}$, or two of the angles tend to 0 and the third to π, as $t \to 0$. This argument involves the choice of a sequence of values of t such that the corresponding $Y_k(t)$, $k = 1, 2, 3$, tend to finite limits. If we consider another sequence of values of t tending to 0 such that the corresponding sequences of values of $Y_k(t)$ also converge to finite limits, then it may happen that the above alternatives get interchanged. That is to say, the points represented by $(X_k(t), Y_k(t))$ may tend to the vertices of an equi-

lateral triangle as $t \to 0$ through one sequence of values, while they way tend to collinear points through another sequence. But, since the angles of the triangle are determined by the sides $r^*_{kl}(t)$ which are continuous functions of t in $0 < t \leq \tau$, and are bounded away from zero as $t \to 0$, we see that the angles are also continuous function of t in $0 < t \leq \tau$. Hence the above possibility canont happen and we conclude that the limiting positions of the points in the plane determined by (x^*_k, y^*_k), $k = 1, 2, 3$, are either at the vertices of an equilateral triangle or in three collinear points. This completes the proof of Theorem 3.2.2. □

Theorem 3.2.3. *If the limiting configuration of the big triangle is an equilateral triangle, then the side of the triangle is given by the positive cube root of*

$$r^3 = \frac{9}{2}(m_1 + m_2 + m_3). \tag{3.2.23}$$

Proof. If the sides of the big triangle are $R_{kl}(t)$, $k \neq l$, then $R_{kl}(t) \to$ **129**
$R_{kl} = r$, $k \neq l$, $k, l = 1, 2, 3$. Once again denoting by X_k the limiting values of $X_k(t)$ as $t \to 0$, we have from (3.2.21) the following algebraic equation satisfied by X_1, X_2, X_3:

$$\begin{aligned}
-\frac{2}{9}X_1 &= \frac{m_2}{r^3}(X_2 - X_1) + \frac{m_3}{r^3}(X_3 - -X_1),\\
-\frac{2}{9}X_2 &= \frac{m_1}{r^3}(X_1 - X_2) + \frac{m_3}{r^3}(x_3 - X_2),\\
-\frac{2}{9}X_3 &= \frac{m_1}{r^3}(X_1 - X_3) + \frac{m_2}{r^3}(X_2 - X_3).
\end{aligned}$$

Since $m_1X_1 + m_2X_2 + m_3X_3 = 0$, we obtain from these equations

$$-\frac{2}{9}X_k = -(m_1 + m_2 + m_3)X_k r^{-3},\ k = 1, 2, 3.$$

Similarly we have for Y_k,

$$-\frac{2}{9}Y_k = -(m_1 + m_2 + m_3)Y_k r^{-3},\ k = 1, 2, 3.$$

Since by (3.2.19), $(R_{kl}(t))^{-1} = (r^*_{kl}(t))^{-1} = 0(1)$ as $t \to 0$, it follows that $R_{kl}(t)$ is bounded away from zero. Thus at least one of the X_k, Y_k,

$k = 1, 2, 3$, is different from zero and we obtain

$$r^3 = \frac{9}{2}(m_1 + m_2 + m_3),$$

which proves the assertion.

Theorem 3.2.3 can also be used to determine explicitly the constant $\kappa > 0$ in the asymptotic estimates for $q^*(t)$, $\dot{q}^*(t), \ldots$ in the equilateral case. Since $U \sim \frac{2}{9}\kappa t^{-2/3}$ as $t \to 0$ and $U^*(q^*)t^{-2/3} = U(q)$, we have

$$U^*(q^*) \sim \frac{2}{9}\kappa \text{ as } t \to 0. \tag{3.2.24}$$

130 The distances $R_{kl}(t)$ are invariant under orthogonal transformations an since $R_{12} = R_{23} = R_{13} = r$, it follows that as $t \to 0$,

$$U^* \to \frac{m_1m_2 + m_2m_3 + m_1m_3}{r} \tag{3.2.25}$$

(3.2.24) and (3.2.25) together imply that

$$\frac{2}{9}\kappa = (m_1m_2 + m_2m_3 + m_1m_3)r^{-1},$$

which determines κ explicitly in terms of the masses.

Next we consider the collinear case. Then the limiting distances R_{12}, R_{23}, R_{31} are no longer equal. Let $\rho = \max(R_{12}, R_{23}, R_{31})$. Suppose that P_1 and P_3 are at the distance ρ at $t = 0$: $\rho^2 = (X_3 - X_1)^2 + (Y_3 - Y_1)^2$. P_2 lies between P_1 and P_3; let $R_{23} = \omega\rho$ where $0 < \omega < 1$. Then $R_{12} = (1 - \omega)\rho$. So

$$R_{31} = \rho, R_{23} = \omega\rho,\ R_{12} = (1 - \omega)\rho. \tag{3.2.26}$$

Once again we make use of the equations (3.2.21). Since the centre of gravity remains fixed at the origin, it follows that the equations satisfied by X_1, X_2, X_3 are not linearly independent. We obtain as in Theorem 3.2.3,

$$\frac{2}{9}(X_1 - X_3) = \frac{2}{9}\rho = \frac{m_1}{\rho^2} + \frac{m_2}{\omega^2\rho^2} + \frac{m_2}{(1-\omega)^2\rho^2} + \frac{m_3}{\rho^2},$$

$$\frac{2}{9}(X_2 - X_3) = \frac{2}{9}\omega\rho = \frac{m_1}{\rho^2} - \frac{m_1}{(1-\omega)^2\rho^2} + \frac{m_2}{\omega^2\rho^2} + \frac{m_3}{\omega^2\rho^2} \tag{3.2.27}$$

or
$$\begin{cases} \frac{2}{9}\rho^3 = m_1 + m_2(\omega^{-2} + (1-\omega)^2) + m_3 \\ \frac{2}{9}\omega\rho^3 = m_1(1-(1-\omega)^{-2}) + m_2\omega^{-2} + m_3\omega^{-2}. \end{cases} \tag{3.2.28}$$

Eliminating ρ between (3.2.27) and (3.2.28) we get **131**

$$m_1(1-\omega-(1-\omega)^{-2}) + m_2((1-\omega)\omega^{-2} - (1-\omega)^{-2}\omega) + m_2(\omega^{-2}-\omega) = 0.$$

Hence ω satisfies the equation

$$m_1((1-\omega)^3-1)\omega^2 + m_2((1-\omega)^3-\omega^3) + m_2((1-\omega)^2 - \omega^3(1-\omega)^2) = 0, \tag{3.2.29}$$

which is an algebraic equation of the fifth degree. This equation has only one root ω in $0 < \omega < 1$. This can be seen as follows: we can write (3.2.29) in the form

$$\frac{m_1 + m_2\omega}{m_1 + m_2\omega^{-2}} = \frac{m_3 + m_2(1-\omega)}{m_3 + m_2(1-\omega)^{-2}}. \tag{3.2.30}$$

Both sides of (3.2.30) are continuous functions of ω in $0 < \omega < 1$. As ω increases from 0 to 1, the left side of (3.2.30) increases from 0 to 1, while the right side decreases from 1 to 0. Hence there exists just one real number ω in $0 < \omega < 1$ satisfying (3.2.30). This unique root is completely determined by the masses m_1, m_2, m_3. Substituting this value of ω in (3.2.27) we obtain ρ as the positive cube root and hence we obtain also the constant κ explicitly in the asymptotic relation $U^* \sim \frac{2}{9}\kappa$ as $t \to 0$, since

$$U^* \to \frac{m_1 m_2}{(1-\omega)\rho} + \frac{m_2 m_3}{\omega\rho} + \frac{m_1 m_3}{\rho}.$$

This corresponds to the case in which P_2 lies between P_1 and P_3. The **132**
other two possibilities are obtained by cyclic permutation of (1,2,3).

Thus, if there is a general collision, we get as the limiting configuration either an equilateral triangle or three collinear points. It is a remarkable fact that the fifth degree equation (3.2.29) for ω already appears in the work of Euler (1767). Euler considered the one-dimensional problem in which all the points are collinear and studied the particular solution giving a general collision in this case. □

3 A particular solution

While we have proved that the limiting configuration of the big triangle is either an equilateral triangle or a set of three collinear points, we have not as yet proved that the triangle itself has a *limiting position* relative to the old fixed coordinate system. We do not also know yet whether the two limiting possibilities can actually be realized while remaining in a fixed coordinate system. We shall now show that the two cases, the equilateral case and the collinear case, do in fact occur. This will be done by giving an explicit particular solution of the three-body problem near $t = 0$, the time of a general collision.

133 We consider the case in which

$$q(t) = q^* \cdot g(t),\ q = x_k, y_k,\ k = 1, 2, 3, \tag{3.3.1}$$

where the q^* are unknown constants, not all zero, and $g(t)$ is an unknown twice continuously differentiable function of t in the interval $0 < t \leq \tau$, and since $q(t)$ should tend to zero as $t \to 0$, we assume that $g(t) \to 0$ as $t \to 0$. Then the differential equations of motion take the form

$$mq^*\ddot{g} = U^*_{q^*} g^{-2},\ \text{or, } mq^*\ddot{g}g^2 = U^*_{q^*}, \tag{3.3.2}$$

where $U^*(q^*) = U(q^*)$, so that by the homogeneity of U, we have $U_q = U^*_{q^*} g^{-2}$. Since q^* are constants, not all zero, the right sides in (3.3.2) are constants and hence

$$\ddot{g}g^2 = \frac{1}{mq^*} U^*_{q^*}$$

is a constant (we take a $q^* \neq 0$). This constant cannot be zero as otherwise $U^*_{q^*} = 0$ for all coordinates q and from the relation $\sum_{q^*} U^*_{q^*} q^* = -U^*$, it would follow that U^* is the constant 0 and this is not the case, by definition of U^*. In view of the considerations of the last section, we take this constant to be $-\dfrac{2}{9}$ so that we have

$$\ddot{g}g^2 = -\frac{2}{9}, \tag{3.3.3}$$

and hence the system of equations (3.3.2) becomes

$$-\frac{2}{9}q^* = \frac{1}{m}U^*_{q^*},$$

which is of the same form as the system of algebraic equations (3.2.21) 134
satisfied by the variables q^* in § 2. Since the equations of motion are invariant under an orthogonal transformation in the plane of the motion, we can take the x-axis to be parallel to the direction of the vector determined by the points (x_3^*, y_3^*) and (x_1^*, y_1^*) and passing through the centre of gravity, which is assumed fixed at the origin. Hence in the new coordinate system $Y_1 = Y_3$ and one can show, using the argument given earlier, that the points defined by (x_k^*, y_k^*), $k = 1, 2, 3$, are either collinear or at the vertices of an equilateral triangle. Hence in order to prove that the two limiting possibilities occur, we shall determine the function $g(t)$ explicitly.

First of all, we observe that the function g cannot vanish anywhere in the interval $0 < t \leq \tau$ by (3.3.3). Integrating the equation $2\dot{g}\ddot{g} = -\dfrac{4}{9}\dot{g}g^{-2}$ we obtain

$$\dot{g}^2 = \frac{4}{9}\left(\frac{1}{g} + C\right),$$

where C is a constant of integration. From this we get

$$\frac{3}{2}\frac{\dot{g}}{\sqrt{C + g^{-1}}} = 1, \text{ or, } \frac{3}{2}\frac{\dot{g}\sqrt{g}}{\sqrt{1 + Cg}} = 1,$$

and this, on integration from 0 to g using the fact that $g(t) \to 0$ as $t \to 0$, gives

$$\frac{3}{2}\int_0^g \frac{\sqrt{g}}{\sqrt{1 + Cg}}\,dg = t.$$

Once again, as $g(t) \to 0$, it follows that $1 + Cg \to 1$ and hence, for t 135 sufficiently small in $0 < t \leq \tau$, $Cg(t)$ is also small. Then we can expand $(1 + Cg)^{\frac{1}{2}}$ as a power-series in $g(t)$ (with parameter C) which converges for sufficiently small t and we have

$\frac{3}{2}\int\limits_0^g \sqrt{g}$ (1 + power-series in g without constant term) $dg = t$, that is, since $g(t) \to 0$ as $t \to 0$,

$g^{3/2}$ (1 + power-series in g without constant term) $= t$. Hence by inversion we have

$$g = t^{2/3} + \text{ power-series in } t^{2/3} \text{ with term of degree } \geq 2, \tag{3.3.4}$$

and the power-series converges for t sufficiently small. The integration above could also be carried out by using trigonometric or hyperbolic functions (according as C is negative or positive) and we could obtain $g(t)$.

The simplest solution for $g(t)$ is the one corresponding to $C = 0$. In this case $\frac{3}{2}\sqrt{g}\dot{g} = 1$ and on integration, $g(t) = t^{2/3}$ and hence we have $q = q^* t^{2/3}$. This proves that both the alternatives can occur in this special case. We have already mentioned that the one-dimensional problem in which the three points situated on a straight line have a general collision on the line was treated by Euler in 1767.

We now proceed to determine the constants q^* of (3.3.1) in some cases. We have assumed throughout that the centre of gravity remains
136 fixed at the origin and so $\sum\limits_{k=1}^{3} m_k x_k^* = 0 = \sum\limits_{k=1}^{3} m_k y_k^*$. We apply an orthogonal transformation in the plane of motion and choose the X-axis to be parallel to the direction of the vector defined by (x_3^*, y_3^*) and (x_1^*, y_1^*) and passing through the centre of gravity. Hence in the new coordinates X, Y, we have $Y_1 = Y_3$ and $X_1 \geq X_3$, where (X_k, Y_k) are the new coordinate of the points (x_k^*, y_k^*), $k = 1, 2, 3$. First consider the equilateral case. Then we have proved that the side of the equilateral triangle is given by $\frac{2}{9}r^3 = m_1 + m_2 + m_3$. Let $m = m_1 + m_2 + m_3$. Since the centre of gravity is at the origin,

$$\sum_{k=1}^{3} m_k X_k = 0, \ \sum_{k=1}^{3} m_k Y_k = 0. \tag{3.3.5}$$

Since the triangle is equilateral, we have

$$X_1 - X_3 = r,\ X_2 - X_3 = \frac{1}{2}r,\ Y_1 - Y_3 = 0,\ Y_2 - Y_3 = \frac{r}{2}\sqrt{3}. \quad (3.3.6)$$

Using (3.3.5) we get

$$\sum_{k=1}^{3} m_k X_k = m_1(X_1 - X_3) + m_2(X_2 - X_3) + mX_3 = m_1 r + \frac{1}{2}m_2 r + mX_3 = 0,$$

$$\sum_{k=1}^{3} m_k Y_k = mY_1 - (m_2 + m_3)Y_1 + m_2 Y_2 + m_3 Y_3 = mY_1 + \frac{1}{2}m_2 r\sqrt{3} = 0.$$

From these and (3.3.6) we obtain

$$X_3 = -\frac{m_1 + \frac{1}{2}m_2}{m}r,\ X_1 = \frac{\frac{1}{2}m_2 + m_3}{m_1}r,\ X_2 = \frac{1}{2}\frac{m_3 - m_1}{m}r, \quad (3.3.7)$$

$$Y_1 = Y_3 = -\frac{1}{2}\frac{m_2}{m}r\sqrt{3},\ Y_2 = \frac{1}{2}\frac{m_1 + m_3}{m}r\sqrt{3}. \quad (3.3.8)$$

Thus, if $g(t) = t^{2/3}$, the original coordinates are given by **137**

$$X_k t^{2/3},\ Y_k t^{2/3},\ k = 1, 2, 3. \quad (3.3.9)$$

Differentiating with respect to t we get

$$(X_2 t^{2/3})^{\cdot} = \frac{1}{3}\frac{m_3 - m_1}{m}rt^{-1/3}, (Y_2 t^{2/3})^{\cdot} = \frac{1}{\sqrt{3}}\frac{m_1 + m_3}{3}rt^{-1/3}. \quad (3.3.10)$$

In the collinear case, we have $Y_1 = Y_2 = Y_3$ and since $\sum_{k=1}^{3} m_k X_k = 0$ and $X_1 - X_3 = \rho$, $X_2 - X_3 = \omega\rho$, we get $m_1\rho + m_2\omega\rho + mX_3 = 0$, so that

$$X_3 = -\frac{m_1 + m_2\omega}{m}\rho,\ X_1 = \frac{m_2(1-\omega) + m_3}{m}\rho,\ X_2 = \frac{m_3\omega - m_1(1-\omega)}{m}\rho. \quad (3.3.11)$$

The originale coordinates in this case are

$$X_k t^{2/3},\ Y_k t^{2/3} = 0,\ k = 1, 2, 3.$$

4 Reduction to a rotating coordinate system

We shall now go back to the general problem of collision. We have already exhibited particular solutions which involve one parameter (namely, the constant C) to show that both the alternatives can be realized for the limiting configuration of the big triangle. The solutions

$$q = q^* g = q^* t^{2/3} + \ldots, q^* = X_k, Y_k, K = 1, 2, 3,$$

suggests that in the general case we may expect to get for $q = q(t)$
138 power-series in the variable $t^{1/3}$ starting with the term $t^{2/3}$. However, this is not true; the general solution involves many more parameters. The difficulty of the problem consists in the fact that we cannot yet prove (this will be proved only at the end) that the big triangle referred to a *fixed* coordinate system has a limiting position as $t \to 0$; all that we have proved so far is the existence of a limiting configuration relative to a rotating coordinate system. The triangle itself may go on rotating above its centre of gravity, assumed fixed at the origin, because in our proof we have made use of an orthogonal transformation in the plane of motion dependeing on the time variable t. We cannot yet determine the limiting position of the big triangle. In order to study this problem more closely we proceed as follows.

We use a fixed coordinate system relative to the *initial* position of the big triangle. We shall first reduce the system of differential equations of motion to one containing a smaller number of equations. The idea is to introduce relative coordinates of P_1 and P_2 with respect to P_3 as we did in the case of simple collisions in Chapter 2 and to make use of the general theory of transformations.

Let (x_k, y_k), $k = 1, 2, 3$, be the coordinates of P_k at time t with respect to the fixed coordinate system through the origin. Let the relative coordinates of P_1 and P_2 with respect to P_3 be (ξ_1, ξ_2) and (ξ_3, ξ_4) respectively; that is,

$$\xi_1 = x_1 - x_3,\ \xi_2 = y_1 - y_3,\ \xi_3 = x_2 - x_3,\ \xi_4 = y_2 - y_3. \tag{3.4.1}$$

139 Since the centre of gravity remains fixed at the origin, we have $m_1\xi_1 +$

$m_2\xi_3 + mX_3 = 0$, $m_1\xi_2 + m_2\xi_4 + mY_3 = 0$ where $m = m_1 + m_2 + m_3$. Then we have

$$
\begin{aligned}
x_3 &= -\frac{m_1}{m}\xi_1 - \frac{m_2}{m}\xi_3,\\
y_3 &= -\frac{m_1}{m}\xi_2 - \frac{m_2}{m}\xi_4,\\
x_1 &= \xi_1 + x_3 = \frac{m_2 + m_3}{m}\xi_1 - \frac{m_2}{m}\xi_3,\\
x_2 &= \xi_3 + x_3 = -\frac{m_1}{m}\xi_1 + \frac{m_1 + m_3}{m}\xi_3,\\
y_1 &= \xi_2 + y_3 = \frac{m_2 + m_3}{m}\xi_2 - \frac{m_2}{m}\xi_4,\\
y_2 &= \xi_4 + y_3 = -\frac{m_1}{m}\xi_2 + \frac{m_1 + m_3}{m}\xi_4.
\end{aligned}
\tag{3.4.2}
$$

If we set

$$\eta_1 = m_1\dot{x}_1,\ \eta_2 = m_1\dot{y}_1,\ \eta_3 = m_2\dot{x}_2,\ \eta_4 = m_2\dot{y}_2, \tag{3.4.3}$$

then, since the centre of gravity remains fixed at the origin,

$$
\begin{aligned}
m_3\dot{x}_3 &= -m_1\dot{x}_1 - m_2\dot{x}_2 = -(\eta_1 + \eta_3),\\
m_3\dot{y}_3 &= -m_1\dot{y}_1 - m_2\dot{y}_2 = -(\eta_2 + \eta_4).
\end{aligned}
\tag{3.4.4}
$$

Since $r_{13}^2 = \xi_1^2 + \xi_2^2$, $r_{23}^2 = \xi_3^2 + \xi_4^2$, $r_{12}^2 = (\xi_1 - \xi_3)^2 + (\xi_2 - \xi_4)^2$, it follows that the potential function U is now a function of the variables $\xi_1, \ldots, \xi_4$ alone. On the other hand,

$$
\begin{aligned}
T &= \frac{1}{2}\sum_{k=1}^{3} m_k(\dot{x}_k^2 + \dot{y}_k^2)\\
&= \frac{1}{2}\left(\frac{1}{m_1} + \frac{1}{m_3}\right)\left(\eta_1^2 + \eta_2^2\right) + \frac{1}{2}\left(\frac{1}{m_2} + \frac{1}{m_3}\right)\left(\eta_3^2 + \eta_4^2\right)\\
&\qquad + \frac{1}{m_3}\left(\eta_1\eta_3 + \eta_2 + \eta_4\right).
\end{aligned}
$$

If E denotes the total energy $T - U$, then the equations of motion can be **140**
written as a Hamiltonian system of eight equations:

$$\dot{\xi}_k = E_{\eta_k},\ \dot{\eta}_k = -E_{\xi_k}, k = 1, \ldots, 4. \tag{3.4.5}$$

We observe that E does not depend on the variable t explicitly. If there is a general collision at $t = 0$, we have already seen that $x_k, y_k = 0(t^{2/3})$ and $\dot{x}_k, \dot{y}_k = 0(t^{-1/3})$, $k = 1, \ldots, 4$, as $t \to 0$. Since the ξ_k are linear functions of x_k and y_k, and the η_k linear functions of $\dot{x}_k$ and $\dot{y}_k$, it follows that, as $t \to 0$,

$$\xi_k = 0(t^{2/3}),\ \eta_k = 0(t^{-1/3}),\ k = 1, \ldots, 4. \tag{3.4.6}$$

If we set $x_k = x_k^* t^{2/3}$, $y_k = y_k^* t^{2/3}$, $\xi_k = \xi_k^* t^{2/3}$, $\eta_k = \eta_k^* t^{-1/3}$, then we see that $x_k^*, y_k^*, \xi_k^*, \eta_k^*$ are $0(1)$ as $t \to 0$.

Now we introduce a rotating coordinate system with origin at P_3 and x-axis along the direction of the vector P_3P_1, i.e. at any instant t we translate the origin to P_3 with the direction of the x, y-axes in the plane of the triangle preserved. We want to consider the limiting *position* of the big triangle with respect to the fixed coordinate system. For this purpose, suppose that at time $t, 0 < t \leq \tau$, the vector P_3P_1 makes an angle $p_4 = p_4(t)$ (positively oriented) with the direction of the x-axis in the old fixed coordinate system. The main difficulty is to obtain the behaviour of p_4 as $t \to 0$. The introduction of the new rotating coordinate system means a transformation of the variables (ξ, η) into new variables which can be described as follows. Let us set

$$c = \cos p_4,\ s = \sin p_4, \tag{3.4.7}$$

141 and let $(p_1, 0)$, (p_2, p_3) denote the coordinates of P_1, P_2 respectively in the new coordinate system; P_3 is $(0, 0)$. Here $p_k = p_k(t)$, $k = 1, 2, 3$. Then the relative coordinates (ξ_1, ξ_2) and (ξ_3, ξ_4) of P_1 and P_2 in the old system are given by

$$\xi_1 = p_1 c - o \cdot s,\ \xi_2 = p_1 s + oc,\ \xi_3 = p_2 c - p_3 s,\ \xi_4 = p_2 s + p_3 c. \tag{3.4.8}$$

The equations (3.4.8) define a transformation of the variables $\xi_1, \ldots, \xi_4$ to $p_1, \ldots, p_4$, and we claim that this can be extended to a canonical transformation of the eight independent variables $\xi_1, \ldots, \xi_4$, $\eta_1, \ldots, \eta_4$. The extension can be done by means of the generating function

$$W = \eta_1 p_1 c + \eta_2 p_1 s + \eta_3 (p_2 c - p_3 s) + \eta_4 (p_2 s + p_3 c). \tag{3.4.9}$$

It is clear that W is linear in η_k and $|W_{p_k\eta_l}| = p_1 \neq 0$. By our general theory (Chapter 1, § 2), the full transformation is obtained by setting $W_{\eta_k} = \xi_k$, $W_{p_k} = q_k$, $k = 1, \ldots, 4$. It is clear from the definition of W that the first set of conditions is satisfied and hence the canonical transformation thus obtained extends (3.4.8). Then we have

$$q_1 = \eta_1 c + \eta_2 s,\ q_2 = \eta_3 c + \eta_4 s,\ q_3 = -\eta_3 s + \eta_4 c,$$
$$q_4 = p_1(-\eta_1 s + \eta_2 c) + p_2(-\eta_2 s + \eta_4 c) - p_3(\eta_3 c + \eta_4 s). \qquad (3.4.10)$$

We shall introduce an auxiliary variable q_o defined by **142**

$$q_o = -\eta_1 s + \eta_2 c, \qquad (3.4.11)$$

and then we can write

$$q_4 = p_1 q_o + p_2 q_3 - p_3 q_2. \qquad (3.4.12)$$

The last equation can be solved for q_o since $p_1 \neq 0$, and we can write $q_o = (q_4 - p_2 q_3 + p_3 q_2) p_1^{-1}$. We could also take this as the definition of q_o. We can express $\eta_1, \ldots, \eta_4$ in terms of $q_o, \ldots, q_3$, using (3.4.10), (3.4.11) and (3.4.12):

$$\eta_1 = q_1 c - q_o s,\ \eta_2 = q_1 s + q_o c,\ \eta_3 = q_2 c - q_3 s,\ \eta_4 = q_2 s + q_3 c. \quad (3.4.13)$$

Then we have $r_{13} = (\xi_1^2 + \xi_2^2)^{1/2} = p_1$, $r_{23} = (\xi_3^2 + \xi_4^2)^{1/2} = (p_2^2 + p_3^2)^{1/2}$, and $r_{12} = ((\xi_1 - \xi_3)^2 + (\xi_2 - \xi_4)^2)^{1/2} = ((p_1 - p_2)^2 + p_3^2)^{1/2}$. So we can express U as a function of p_1, p_2, p_3 alone. Also

$$T = \frac{1}{2}\left(\frac{1}{m_1} + \frac{1}{m_3}\right)\left(q_0^2 + q_1^2\right) + \frac{1}{2}\left(\frac{1}{m_2} + \frac{1}{m_3}\right)\left(q_2^2 + q_3^2\right) + \frac{1}{m_3}\left(q_1 q_2 + q_o q_3\right).$$

This shows that the total energy $T - U$ expressed in terms of the new variables (p, q) does not contain p_4. We already know the asymptotic behaviour of p_1, p_2, p_3; also $q_o, \ldots, q_3$ behave nicely. We do not know the behaviour of p_4, but this bad coordinate disappears from the function E.

Since W does not contain the variable t explicitly, we have $E(\xi, \eta) = \mathbb{E}(p, q)$ and we denote $\mathbb{E}(p, q)$ by $E(p, q)$ itself.

The Hamiltonian system (3.3.5) now becomes 143

$$\dot{p}_k = E_{q_k}, \dot{q}_k = -E_{p_k}, k = 1, 2, 3, \dot{p}_4 = E_{q_4}, \dot{q}_4 = 0. \tag{3.4.14}$$

Hence q_4 is a constant. We shall now prove that when there is a general collision, this constant has necessarily to be zero.

Theorem 3.4.1. *If there is a general collision at $t = 0$, then $q_4(t) \equiv 0$.*

Proof. By Sundman's theorem (Theorem 2.2.2), the constants of angular momenta λ, μ, ν, all vanish. In particular, $\lambda = \sum_{k=1}^{3} m_k(x_k\dot{y}_k - y_k\dot{x}_k) = 0$. q_4 is given by (3.4.12). But from (3.4.8), and (3.4.13) $\xi_1\eta_2 - \xi_2\eta_1 = p_1q_o$, $\xi_3\eta_4 - \xi_4\eta_3 = p_2q_3 - p_3q_2$, so that we can write

$$q_4 = (\xi_1\eta_2 - \xi_2\eta_1) + (\xi_3\eta_4 - \xi_4\eta_3).$$

Moreover, from the definitions (3.4.1) and (3.4.3) of ξ_k, η_k, we have

$$\begin{aligned}
\xi_1\eta_2 - \xi_2\eta_1 &= m_1(x_1 - x_3)\dot{y}_1 - m_1(y_1 - y_3)\dot{x}_1 \\
&= m_1(x_1\dot{y}_1 - y_1\dot{x}_1) - m_1(x_3\dot{y}_1 - y_3\dot{x}_1), \\
\xi_3\eta_4 - \xi_4\eta_3 &= m_2(x_2 - x_3)\dot{y}_2 - m_2(y_2 - y_3)\dot{x}_2 \\
&= m_2(x_2\dot{y}_2 - y_2\dot{x}_2) - m_2(x_3\dot{y}_2 - y_3\dot{x}_2).
\end{aligned}$$

Hence we have

$$q_4 = \sum_{k=1}^{2} m_k(x_k\dot{y}_k - y_k\dot{x}_k) - (m_1\dot{y}_1 + m_2\dot{y}_2)x_3 + (m_1\dot{x}_1 + m_2\dot{x}_2)y_3.$$

144 Since the centre of gravity remains fixed at the origin of the coordinate system (x, y), we have $m_1\dot{x}_1 + m_2\dot{x}_2 = -m_3\dot{x}_3$, $m_1\dot{y}_1 + m_2\dot{y}_2 = -m_3\dot{y}_3$ and so,

$$q_4 = \sum_{k=1}^{3} m_k(x_k\dot{y}_k - y_k\dot{x}_k) = \lambda,$$

and we know that $\lambda = 0$. This completes the proof of the theorem.

Thus the hamiltonian system (3.4.14) takes the form

$$\dot{p}_k = (E_{qk})_{q_4=0}, \dot{q}_k = -(E_{p_k})_{q_4=0},\ k = 1, 2, 3,\ \dot{p}_4 = (E_{q_4})_{q_4=0}. \tag{3.4.15}$$

We shall now introduce the variables p_k^* by setting

$$p_k = p_k^* t^{2/3}, \ k = 1, 2, 3. \tag{3.4.16}$$

Since we know by (3.4.1) and (3.4.8) that

$$\begin{aligned} p_1 &= \xi_1 c + \xi_2 s = (x_1 - x_3)c + (y_1 - y_3)s, \\ p_2 &= \xi_3 c + \xi_4 s = (x_2 - x_3)c + (y_2 - y_3)s, \\ p_3 &= -\xi_3 s + \xi_4 c = -(x_2 - x_3)s + (y_2 - y_3)c, \end{aligned}$$

and (x_k^*, y_k^*), $k = 1, 2, 3$, are the vertices of the big triangle with respect to the original system of coordinates with origin at the centre of gravity, we can take the coordinates of these vertices in the new coordinate system with origin at P_3 and x-axis parallel to P_3P_1, to be $(p_1^*, 0)$, (p_2^*, p_3^*), $(0, 0)$. We know by Theorem 3.2.2, that the big triangle has a limiting configuration which is either an equilateral triangle or a set of three collinear points. (In the latter case, of course, there are three possibilities, but we restrict ourselves to one of them, namely the case in which P_2 is between P_1 and P_3; the two other cases are similar). In other words, we have proved that relative to the rotating coordinate system, $p_k^*, k = 1, 2, 3$, tends to a finite limit as $t \to 0$. We shall denote these limits by $\bar{p}_1, \bar{p}_2$ and $\bar{p}_3$. We have also determined the limiting mutual distances, in fact, in the equilateral case we have **145**

$$\bar{p}_1 = r, \ \bar{p}_2 = \frac{1}{2} r, \ \bar{p}_3 = \frac{1}{2}\sqrt{3} r, \tag{3.4.17}$$

and in the collinear case,

$$\bar{p}_1 = \rho, \ \bar{p}_2 = \omega\rho, \bar{p}_3 = 0. \tag{3.4.18}$$

Here r, ρ, ω are given by (3.2.23), (3.2.27) and (3.2.30); they are uniquely determined by the masses. On the other hand, we know that since $\dot{x}_k, \dot{y}_k, (k = 1, 2, 3)$, are $0(t^{-1/3})$ as $t \to 0$, also $\dot{\eta}_k = 0(t^{-1/3})$ as $t \to 0$, $k = 1, 2, 3, 4$. Hence $q_k, k = 0, \ldots, 3$, given by (3.4.10), being linear in $\eta_1, \ldots, \eta_4$, are also $0(t^{-1/3})$ as $t \to 0$, and moreover $q_4 \equiv 0$. If we introduce the new variables $q_k^*, k = 0, \ldots, 3$, by setting

$$q_k = q_k^* t^{-1/3}, \tag{3.4.19}$$

then $q_k^* = 0(1)$ as $t \to 0$. We would expect that q_k^* tend to finite limits as $t \to 0$. We show that this is in fact the case. □

146 **Theorem 3.4.2.** *If there is a genral collision at $t = 0$, then q_k^* tend to finite limits $\bar{q}_k$ as $t \to 0$.*

Proof. We consider the function $\sigma = \sum_{k=1}^{3} m_k(x_k^2 + y_k^2)$ which, on differentiation with respect to t, gives $\frac{1}{2}\dot{\sigma} = \sum_{k=1}^{3} m_k(x_k\dot{x}_k + y_k\dot{y}_k)$. This gives, in view of the fact that the centre of gravity remains fixed at the origin and so $m_3\dot{x}_3 = -(m_1\dot{x}_1 + m_2\dot{x}_2)$ and $m_3\dot{y}_3 = -(m_1\dot{y}_1 + m_2\dot{y}_2)$,

$$\begin{aligned}\frac{1}{2}\dot{\sigma} &= \sum_{k=1}^{2} m_k((x_k\dot{x}_k + y_k\dot{y}_k) - x_3(m_1\dot{x}_1 + m_2\dot{x}_2) - y_3(m_1\dot{y}_1 + m_2\dot{y}_2)) \\ &= \sum_{k=1}^{2} m_k((x_k - x_3)\dot{x}_k + (y_k - y_3)\dot{y}_k).\end{aligned}$$

Using (3.4.1), (3.4.3), (3.4.8) and (3.4.13) we obtain from this

$$\frac{1}{2}\dot{\sigma} = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3 + \xi_4\eta_4 = p_1q_1 + p_2q_2 + p_3q_3.$$

By Theorem 3.1.2, we have $\dot{\sigma} \sim \frac{4}{3}\kappa t^{1/3}$. If we write $\dot{\sigma}$ in terms of p_k^*, q_k^*, $k = 1, 2, 3$, we get, as $t \to 0$,

$$\frac{1}{2}\dot{\sigma} = (p_1^*q_1^* + p_2^*q_2^* + p_3^*q_3^*)t^{1/3} \sim \frac{2}{3}\kappa t^{1/3}.$$

Hence it follows that, as $t \to 0$,

$$p_1^*q_1^* + p_2^*q_2^* + p_3^*q_3^* \to \frac{2}{3}\kappa. \tag{3.4.20}$$

(The constant κ has already been explicitly determined in terms of the masses in both the equilateral and collinear cases). Since $q_4 = 0$, we also have from (3.4.12), $p_1q_o + p_2q_3 - p_3q_2 = q_4 = 0$ and hence

$$p_1^*q_o^* + p_2^*q_3^* - p_3^*q_2^* = 0. \tag{3.4.21}$$

147

We consider first the equilateral case. By (3.1.3), if there is a general collision at $t = 0$ and u, v are two distinct coordinates among the x_k, y_k, $k = 1, 2, 3$, then as $t \to 0$, $u\dot{v} - v\dot{u} = o(t^{1/3})$, and so in particular,

$$m_1(x_1\dot{y}_1 - y_1\dot{x}_1) = o(t^{1/3}),\ m_2(x_2\dot{y}_2 - y_2\dot{x}_2) = o(t^{1/3}). \qquad (3.4.22)$$

But

$$\begin{aligned} m_1(x_1\dot{y}_1 - y_1\dot{x}_1) &= \left(\frac{m_2+m_3}{m}\xi_1 - \frac{m_2}{m}\xi_3\right)\eta_2 - \left(\frac{m_2+m_3}{m}\xi_2 - \frac{m_2}{m}\xi_4\right)\eta_1 \\ &= \frac{m_2+m_3}{m}(\xi_1\eta_2 - \xi_2\eta_1) - \frac{m_2}{m}(\xi_3\eta_2 - \xi_4\eta_1); \\ m_2(x_2\dot{y}_2 - y_2\dot{x}_2) &= \left(-\frac{m_1}{m}\xi_1 + \frac{m_1 m_3}{m}\xi_3\right)\eta_4 - \left(-\frac{m_1}{m}\xi_2 + \frac{m_1+m_3}{m}\xi_4\right)\eta_3 \\ &= -\frac{m_1}{m}(\xi_1\eta_4 - \xi_2\eta_3) + \frac{m_1+m_3}{m}(\xi_3\eta_4 - \xi_4\eta_3). \end{aligned}$$

Now passing to the variables p_k, q_k using the transformation given by (3.4.8) and (3.4.13), we find from (3.4.22) that, as $t \to 0$,

$$\begin{aligned} &\frac{m_2+m_3}{m}p_1q_o - \frac{m_2}{m}(p_2q_o - p_3q_1) = m_1(x_1\dot{y}_1 - y_1\dot{x}_1) = o(t^{1/3}), \\ &-\frac{m_1}{m}p_1q_3 + \frac{m_1+m_3}{m}(p_2q_3 - p_3q_2) = m_2(x_2\dot{y}_2 - y_2\dot{x}_2) = o(t^{1/3}). \end{aligned}$$

Hence we have the following two additional relations for $p_k^*, k = 1, 2, 3,$ and $q_k^*, k = 0, \ldots, 3$: as $t \to 0$,

$$(m_2+m_3)p_1^*q_o^* - m_2(p_2^*q_o^* - p_3^*q_1^*) = o(1), \qquad (3.4.23)$$

$$-m_1p_1^*q_3^* + (m_1+m_3)\,(p_2^*q_3^* - p_3^*q_2^*) = o(1). \qquad (3.4.24)$$

The four equations (3.4.20), (3.4.21), (3.4.23) and (3.4.24) remain valid 148
as $t \to 0$; since $q^* = 0(1)$ as $t \to 0$ we can therefore write $\overline{p}_k$ in place of p_k^* and obtain the equations in the unknowns $q_o^*, \ldots, q_3^*$: as $t \to 0$,

$$\begin{aligned} \overline{p}_1q_o^* - \overline{p}_3q_2^* + \overline{p}_2q_3^* &= 0, \\ \overline{p}_1q_1^* + \overline{p}_2q_2^* + \overline{p}_3q_3^* &= \frac{2}{3}\kappa + o(1), \\ ((m_2+m_3)\overline{p}_1 - m_2\overline{p}_2)q_o^* + m_2\overline{p}_3q_1^* &= o(1), \end{aligned}$$

$$-(m_1+m_3)\bar{p}_3 q_2^* + ((m_1+m_3)\bar{p}_2 - m_1\bar{p}_1)q_3^* = o(1).$$

Instead of the above asymptotic relations for q^* we shall consider for the moment the associated system of linear equations in q^* with the error terms omitted; we write $\bar{q}$ in place of q^* and we have

$$\begin{aligned}
\bar{p}_1\bar{q}_o - \bar{p}_3\bar{q}_2 + \bar{p}_2\bar{q}_3 &= 0,\\
\bar{p}_1\bar{q}_1 + \bar{p}_2\bar{q}_2 + \bar{p}_3\bar{q}_3 &= \frac{2}{3}\kappa,\\
((m_2+m_3)\bar{p}_1 - m_2\bar{p}_2)\bar{q}_o + m_2\bar{p}_3\bar{q}_1 &= 0,\\
-(m_1+m_3)\bar{p}_3\bar{q}_2 + ((m_1+m_3)\bar{p}_2 - m_1\bar{p}_1)\bar{q}_3 &= 0.
\end{aligned} \tag{3.4.25}$$

149 A solution of this system of linear equations provides a solution of the problem in the special case in which $p_k = \bar{p}_k t^{2/3}$, $q_k = \bar{q}_k t^{-1/3}$. The special solution is given by:

$$\begin{gathered}
p_k = \bar{p}_k t^{2/3}, k = 1, 2, 3;\ p_4 = 0;\ \xi_1 = p_1,\ \xi_2 = 0,\ \xi_3 = p_2, \xi_4 = p_3;\\
x_1 = \frac{m_2+m_3}{m}p_1 - \frac{m_2}{m}p_2,\\
x_2 = -\frac{m_1}{m}p_1 + \frac{m_1+m_3}{m}p_2,\\
y_1 = -\frac{m_2}{m}p_3, y_2 = \frac{m_1+m_3}{m}p_3;\\
q_k = \bar{q}_k t^{-1/3}, k = 0, 1, 2, 3;\\
q_1 = m_1\dot{x}_1 = \frac{m_1}{m}(m_2+m_3)\dot{p}_1 - \frac{m_1 m_2}{m}\dot{p}_2,\\
q_o = m_1\dot{y}_1 = -\frac{m_1 m_2}{m}\dot{p}_3,\\
q_2 = m_2\dot{x}_2 = -\frac{m_1 m_2}{m}\dot{p}_1 + \frac{m_2}{m}(m_1+m_3)\dot{p}_2,\\
q_3 = m_2\dot{y}_2 = \frac{m_2}{m}(m_1+m_3)\dot{p}_3.
\end{gathered}$$

The determinant of the system (3.4.25) of linear equations in $\bar{q}_o, \bar{q}_1, \bar{q}_2, \bar{q}_3$ is seen to be

$$-\bar{p}_1\bar{p}_3(m_1 m_3 \bar{p}_1^2 + m_1 m_2(\bar{p}_1-\bar{p}_2)^2 + m_2 m_3 \bar{p}_2^2 + (m_1+m_3)m_2\bar{p}_3^2). \tag{3.4.26}$$

In the equilateral case, using the values for $\overline{p}_k, k = 1, 2, 3$, given by (3.4.17), this is seen to be $-\frac{1}{2}\sqrt{3}(m_1m_2 + m_2m_3 + m_3m_1)r^4 \neq 0$, and so we obtain the solution of (3.4.25):

$$\begin{aligned}
\overline{q}_o &= -\frac{m_1m_2}{\sqrt{3}m}r,\\
\overline{q}_1 &= \frac{m_1(m_2+2m_3)}{3m}r,\\
\overline{q}_2 &= \frac{m_2(m_3-m_1)r}{3m},\\
\overline{q}_3 &= \frac{m_2(m_1+m_3)r}{\sqrt{3}m}.
\end{aligned} \tag{3.4.27}$$

In the collinear case, since $\overline{p}_1 = \rho$, $\overline{p}_2 = \omega\rho$, $\overline{p}_3 = 0$, it turns out that the determinant (3.4.26) is zero. However, we may, in this case, use instead of the asymptotic relations (3.4.22), the relation $x_1\dot{x}_2 - x_2\dot{x}_1 = o(t^{1/3})$, $y_1\dot{y}_2 - y_2\dot{y}_1 = o(t^{1/3})$ leading to 150

$$m_1m_2(x_1\dot{x}_2 - x_2\dot{x}_1 + y_1\dot{y}_2 - y_2\dot{y}_1) = o(t^{1/3}) \text{ as } t \to 0. \tag{3.4.28}$$

Using the values given above for $x_1, \ldots, \dot{y}_2$, this gives

$$\begin{aligned}
m_1\frac{(m_2+m_3)}{m}p_1q_2 - \frac{m_1m_2}{m}(p_2q_2 + p_3q_3 - p_1q_1)&\\
- m_2\frac{(m_1+m_3)}{m}(p_2q_1 + p_3q_o) = o(t^{1/3})&
\end{aligned}$$

or

$$\begin{aligned}
m_2(m_1+m_3)\overline{p}_3\overline{q}_o - m_2(m_1\overline{p}_1 - (m_1+m_3)\overline{p}_2)\overline{q}_1 + m_1(m_2\overline{p}_2&\\
- (m_2+m_3)\overline{p}_1)\overline{q}_2 + m_1m_2\overline{p}_3\overline{q}_3 = o(1).&
\end{aligned}$$

Now we consider the system (3.4.25) with the fourth equation replaced by this, without the error term on the right. The system is now of rank 4, the determinant is seen to be

$$-\omega(m_2(1-\omega) + m_3)\,(m_1m_2(1-\omega)^2 + m_1m_3 + m_2m_3\omega^2)\rho^4 \neq 0,$$

and we can solve the system as we did earlier and we obtain the solutions

$$
\begin{aligned}
\bar{q}_o &= 0, \bar{q}_1 = \frac{2}{3} m_1 \frac{m_2(1-\omega)+m_3}{m}\rho,\\
\bar{q}_2 &= \frac{2}{3} m_2 \frac{-m_1(1-\omega)+m_3\omega}{m}\rho,\\
\bar{q}_3 &= 0. \qquad (3.4.29)
\end{aligned}
$$

We have thus proved the theorem in the special case of the particular solution $p_k = p_k^* t^{2/3}$, $q_k = \bar{q}_k t^{-4/3}$, where p_k^* are constants. We now take up the general case. From the relations (3.4.20), (3.4.21), (3.4.23) and (3.4.24) we have the asymptotic equations satisfied by $q_o^*, \ldots, q_3^*$: as $t \to 0$,

$$
\begin{aligned}
p_1^* q_o^* + p_2^* q_3^* - p_3^* q_2^* &= 0,\\
p_1^* q_1^* + p_2^* q_2^* + p_3^* q_3^* &= \frac{2}{3}\kappa + o(1),\\
((m_2+m_3)p_1^* - m_2 p_2^*) q_o^* + m_2 p_3^* q_1^* &= o(1),\\
-m_1 p_1^* q_3^* + (m_1+m_3)(p_2^* q_3^* - p_3^* q_2^*) &= o(1). \qquad (3.4.30)
\end{aligned}
$$

151 Since $p_k^*(t) \to \bar{p}_k$, $k = 1, 2, 3$, as $t \to 0$, we can write, for t sufficiently near 0, $p_k^*(t) = \bar{p}_k + \epsilon_k(t)$, $\epsilon_k(t) = o(1)$ as $t \to 0$. In the equilateral case, then, recalling the values of $\bar{p}_k$ given by (3.4.17), we can replace p_k^* in the system (3.4.30) by

$$
p_1^* = r + \epsilon_1(t), \;\; p_2^* = \frac{1}{2} r + \epsilon_2(t), \;\; p_3^* = \frac{1}{2} r\sqrt{3} + \epsilon_3(t).
$$

The determinant of the system of linear equations (3.4.30) is a quartic polynomial in p_k^* and hence a continuous function of the variables p_k^*, $k = 1, 2, 3$. Hence, as $t \to 0$, this determinant tends to the determinant of the system (3.4.30) with p_k^* replaced by $\bar{p}_k$. We have seen that the latter determinant is $\frac{\sqrt{3}}{2}(m_1 m_2 + m_2 m_3 + m_3 m_1) r^4 \neq 0$. Hence, for t sufficiently close to 0, the determinant of (3.4.30) is also different from zero, by continuity. Let t be small enough for this condition to hold. Then we consider the system of linear equations

$$
p_1^* q_o^* - p_3^* q_2^* + p_2^* q_3^* = 0
$$

$$\begin{aligned} p_1^* q_1^* + p_2^* q_2^* + p_3^* q_3^* &= \frac{2}{3}\kappa + \delta_1(t), \\ ((m_2 + m_3)p_1^* - m_2 p_2^*)q_o^* + m_2 p_3^* q_1^* &= \delta_2(t) \\ (-m_1 p_1^* + (m_1 + m_3)p_2^*)q_3^* - (m_1 + m_3)p_3^* q_2^* &= \delta_3(t), \end{aligned} \tag{3.4.31}$$

where $\delta_k(t) \to 0$ as $t \to 0$. The determinant of the system (3.4.31) is **152**

$$-\frac{1}{2}\sqrt{3}(m_1m_2 + m_2m_3 + m_3m_1)r^4 + \delta_4(t) \neq 0, \;\; \delta_4(t) \to 0 \text{ as } t \to 0.$$

Hence we can solve the system (3.4.31) and obtain q_k^* as rational functions of p_k^*, δ_k, with the non-vanishing determinant in the denominator. Hence q_k^* tend to finite limits $\bar{q}_k$ as $t \to 0$, and these limits are the same as the solutions of the system (3.4.31) with p_k^* replaced by their limits $\bar{p}_k$ and δ_k replaced by their limit 0. We have already obtained these particular solutions. Hence, in the equilateral case, we find that as $t \to 0$,

$$\begin{gathered} p_1 \sim rt^{2/3}, \;\; p_2 \sim \frac{r}{2}t^{2/3}, p_3 \sim \frac{1}{2}\sqrt{3}rt^{2/3}, \\ q_1 \sim \frac{m_1}{3m}(m_2 + 2m_2)rt^{-1/3}, q_2 \sim \frac{m_2}{3m}(m_3 - m_1)rt^{-1/3}, \\ q_3 \sim \frac{m_2}{\sqrt{3m}}(m_1 + m_3)rt^{-1/3}, q_4 = 0. \end{gathered}$$

Next, in the collinear case, $\bar{p}_1 = \rho$, $\bar{p}_2 = \omega\rho$, $\bar{p}_3 = 0$, where ω, ρ and κ are uniquely determined by the masses. It turns out, as we have seen earlier, that the determinant of the system (3.4.30) tends to zero as $t \to 0$, but we can replace the last of the equations suitably, as we did for the particular solution (namely, by using the relation $m_1m_2(x_1\dot{x}_2 - x_2\dot{x}_1 + y_1\dot{y}_2 - y_2\dot{y}_1) = o(t^{1/3})$ as $t \to 0$), and obtain a linear system in $q_o^*, \ldots, q_3^*$ with determinant $\neq 0$. An argument on the same lines as in the equilateral case applied to the new system now proves that $q_k^*(t)$ tend to finite limits $\bar{q}_k$ as $t \to 0$, $k = 0, \ldots, 3$, and we have

$$\bar{q}_1 = \frac{2m_1}{3m}(m_2(1-\omega) + m_3)\rho, \bar{q}_2 = \frac{2m_2}{3m}(m_3\omega - (1-\omega)m_1)\rho, \bar{q}_3 = 0.$$

Hence we have, in the collinear case, as $t \to 0$, **153**

$$q_1 \sim \frac{2m_1}{3m}(m_2(1-\omega) + m_3)\rho t^{-1/3},$$

$$q_2 \sim \frac{2m_2}{3m}(m_3\omega - (1-\omega)m_1)\rho t^{-1/3}, q_3 = 0.$$

This completes the proof of the theorem. □

We have proved that $q_4 \equiv 0$ for a collision orbit. If we now substitute $q_4 = 0$ in E_{p_k} and E_{q_k}, $k = 1, 2, 3$, then we get a Hamiltonian system

$$\dot{p}_k = (E_{q_k})_{q_4=0},\ \dot{q}_k = -(E_{p_k})_{q_4=0}, k = 1, 2, 3,$$

with six degrees of freedom. If we solve this system for p_k, q_k, $k = 1, 2, 3$, and substitute in the remaining equation $\dot{p}_4 = (E_{q_4})_{q_4=0}$, then we get a differential equation for p_4. We can integrate this to obtain p_4. But we cannot prove that $p_4(t)$ has a limit as $t \to 0$ until we have proved that the integral of the function $(E_{q_4})_{q_4=0}$ converges as $t \to 0$. Hence we cannot determine the behaviour of p_4 as yet. However, in the case of the particular solution $x_k = x_k^* t^{2/3}$, $y_k = y_k^* t^{2/3}$, $k = 1, 2, 3$, where x_k^*, y_k^* are unknown constants, it is clear that p_4 is a constant. By a rotation of the coordinate system we may then assume that $p_4(t) \equiv 0$. For the particular solution with this choice of coordinates, we have $\dot{p}_4 = 0$, i.e. $(E_{q_4})_{q_4=0} = 0$.

154 We had introduced the variables p_k^*, q_k^*, $k = 1, 2, 3$, by setting $p_k = p_k^* t^{2/3}$, $q_k = q_k^* t^{-1/3}$ and we now set formally $p_4^* = p_4$, $q_4^* t^{1/3} = q_4$. With this definition,

$$q_o^* = (q_4^* - p_2^* q_3^* + p_3^* q_2^*)/p_1^*,\ q_o = q_o^* t^{-1/3}.$$

We shall now express the total energy E in terms of the variables p_k^*, q_k^*. If f is a function of the variables p_k, q_k, we shall denote by f^* the function of the variables p_k^*, q_k^* defined by $f^*(p^*, q^*) = f(p^*, q^*)$. Since $p_k = p_k^* t^{2/3}$, $q_k = q_k^* t^{-1/3}$, $k = 1, 2, 3$, and $q_o = q^* o t^{-1/3}$, we see that

$$\begin{aligned} T(q_o, \dots, q_3) &= \frac{1}{2}(\frac{1}{m_1} + \frac{1}{m_3})(q_1^2 + q_0^2) \\ &\qquad + \frac{1}{2}(\frac{1}{m_2} + \frac{1}{m_3})(q_2^2 + q_3^2) + \frac{1}{m_3}(q_1 q_2 + q_o q_3) \\ &= T^*(q_o^*, \dots, q_3^*) t^{-2/3}; \end{aligned}$$

$$U(p_1, p_2, p_3) = \frac{m_1 m_2}{\sqrt{(p_1 - p_2)^2 + p_3^2}} + \frac{m_2 m_3}{\sqrt{p_2^2 + p_3^2}} + \frac{m_1 m_3}{p_1}$$
$$= U^*(p_1^*, p_2^*, p_3^*) t^{-2/3}.$$

So we have

$$E(p_1, p_2, p_3, q_o, \dots, q_3) = E^*(p_1^*, p_2^*, p_3^*, q_o^*, \dots, q_3^*) t^{-2/3}, \qquad (3.4.32)$$

and the differential equations (3.4.15) can be written down in terms of the variables p_k^*, q_k^*:

$$\dot{p}_k = \dot{p}_k^* t^{2/3} + \frac{2}{3} p_k^* t^{-1/3} = E_{q_k} = E^*_{q_k^*} \frac{dq_k^*}{dq_k} t^{-2/3} = E^*_{q_k^*} t^{-1/3}, k = 1, 2, 3;$$
$$\dot{q}_k = \dot{q}_k^* t^{-1/3} - \frac{1}{3} q_k^* t^{-4/3}$$
$$= -E_{p_k} = -E^*_{p_k^*} \frac{dp_k^*}{dp_t} t^{-2/3} = -E^*_{p_k^*} t^{-4/3}, k = 1, 2, 3;$$
$$\dot{p}_4 = \dot{p}_4^* = E^*_{q_4^*} \frac{dq_4^*}{dq_4} t^{-2/3} = E^*_{q_4^*} t^{-1},$$
$$\dot{q}_4 = \dot{q}_4^* t^{1/3} + \frac{1}{3} q_4^* t^{-2/3} = -E^*_{p_4^*} t^{-2/3} (\equiv 0).$$

Finally, then, the equations of motion take the form 155

$$t\dot{p}_k^* = E^*_{q_k^*} - \frac{2}{3} p_k^*; t\dot{q}_k^* = -E^*_{p_k^*} + \frac{1}{3} q^* k, k = 1, 2, 3;$$
$$t\dot{p}_4^* = E^*_{q_4^*}; t\dot{q}_4^* = -\frac{1}{3} q_4^*. \qquad (3.4.33)$$

Thus we have a system of differential equations of the first order in which the right sides are explicitly determined functions of $p_k^*, k = 1, 2, 3$, and $g_k^*, k = 1, \dots, 4$. Now we claim that the right sides of (3.4.33) can be expanded into power-series in the seven independent variables p_k^*, q_k^* in some neighbourhood of $\overline{p}_k, k = 1, 2, 3$, and $\overline{q}_k, k = 1, 2, 3, 4$. In order to see this, we observe that T is a homogeneous polynomial of the second degree in the variables q_k, $k = 0, \dots, 3$ alone and that U is a homogeneous function of degree -1 in $p_k, k = 1, 2, 3$, alone. Then T^* is a homogeneous polynomial of degree 2 in q^*_k and U^* is a homogeneous function of degree -1 in p^*_k. Since $r^*_{12} = ((p^*_1 - p^*_2)^2 + p^{*2}_3)^{1/2}$, $r^*_{23} = (p^{*2}_2 + p^{*2}_3)^{1/2}$, $r^*_{13} = p^*_1$ are the sides of the big triangle and hence do not vanish as $t \to 0$, the denominators in $U^* = m_1 m_2 / r^*_{12} + m_2 m_3 / r^*_{23} + m_1 m_3 / r^*_{13}$ do not vanish as $t \to 0$. Moreover, $p^*_k(t) \to \bar{p}_k$, $k = 1, 2, 3,$, and $q^*_k(t) \to \bar{q}_k$, $k = 1, \ldots, 4$, as $t \to 0$. Hence we can expand U^* as a power-series in p^*_k in a suitable neighbourhood of $\bar{p}_k, k = 1, 2, 3$. Thus $E^* = T^* - U^*$, and hence the right sides in (3.4.33) can be expanded as power-series in all the seven variables in a neighbourhood of $\bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{q}_1, \ldots \bar{q}_4$ which proves our assertion.

156 Since $p^*_k \to \bar{p}_k, q^*_k \to \bar{q}_k$, $k = 1, 2, 3$, and $q^*_4 \to \bar{q}_4$, we can write $p^*_k = \bar{p}_k + \delta_k + \delta_k$, $k = 1, 2, 3$; $q^*_k = \bar{q}_k + \delta_{k+3}$, $k = 1, \ldots 4$ where $\delta_k(t) \to 0$ as $t \to 0$. Then it follows that the right sides of (3.4.33) can be expanded as power-series in the seven independent variables $\delta_1, \ldots, \delta_7$ for $|\delta_k|$ sufficiently small for $k = 1, 2, 3$, and for arbitrary δ_{k+3}, $k = 1, \ldots, 4$. We may also write $p^*_4 = \delta_4$, but this variable does not actually appear on the right sides of (3.4.33).

We now assert that the power-series expansion for the right sides in (3.4.33) in the variables δ_k do not contain constant terms. This can be proved in the following way. We observe that the particular solution $p^*_k = \bar{p}_k t^{2/3}$, $q^*_k = \bar{q}_k t^{-1/3}$, $k = 1, 2, 3$, of the three-body problem satisfies the system of differential equations $\dot{p}_k = E_{q_k}$, $\dot{q}_k = -E_{p_k}$, $k = 1, 2, 3$, and in addition, $p_4 = 0$ by assumption and $q_4 = 0$ for a collision orbit. Then p^*_k, q^*_k also satisfy the differential equations (3.4.33); p^*_k, q^*_k being constants for the particular solution, the left sides of (3.4.33) are zero for this solution and hence the right sides vanish. Thus

$$E^*_{q^*_k}(\bar{p}_k, \bar{q}_k) - \frac{2}{3}\bar{p}_k = 0, \ -E^*_{p^*_k}(\bar{p}_k, \bar{q}_k) + \frac{1}{3}\bar{q}_k = 0,$$
$$E^*_{q^*_4}(\bar{p}_k, \bar{q}_k) = 0, \ -E^*_{p^*_4}(\bar{p}_k, \bar{q}_k) - \frac{1}{3}\bar{q}_4 = 0.$$

But these are precisely the constant terms in the expansions in power-series of the right sides of (3.4.33) and so our assertion is proved.

157 We observe that the variable t appears explicitly on the left side in (3.4.33) and so we shall transform the system into one in a new variable s so that s does not appear explicitly on the left side. For this we

introduce the new variable s by means of the substitution

$$t = e^{-s}, \text{ or } s = \log \frac{1}{t}. \tag{3.4.34}$$

Then $ds = -\dfrac{dt}{t}$. We shall denote the derivative of a function $f(s)$ with respect to s by $f'(s)$. Then

$$\dot{p}_k^* = p_k^{*'} \frac{ds}{dt} = -\frac{1}{t} p_k^{*'}, \text{ or } t\dot{p}_k^* = -p_k^{*'}, \text{ and } t\dot{q}_k^* = -q_k^{*'},$$

so that the system (3.4.33) is transformed into the new system

$$p_k^{*'} = \frac{2}{3} p_k^* - E^*_{q_k^*},\ q_k^{*'} = E^*_{p_k^*} - \frac{1}{3} q_k^*, k = 1,2,3;$$
$$p_4^{*'} = -E^*_{q_4^*}, q_4^{*'} = \frac{1}{3} q_4^*. \tag{3.4.35}$$

Now we introduce the variables $\bar{p}_k + \delta_k, k = 1,2,3$, and $\bar{q}_k + \delta_{k+3}$, $k = 1,\ldots,4$. We have $p_k^{*'} = \delta_k'$, $q_k^{*'} = \delta_{k+3}'$, $k = 1,2,3$, $q_4^{*'} = \delta_7'$, and the right sides of (3.4.35) do not contain constant terms in their power-series expansions and the variable p_4^* does not occur. So the system (3.4.35) takes the form

$$\delta_k' = \sum_{l=1}^{8} a_{kl}\delta_l + \varphi_k(\delta_1,\ldots,\delta_7),\ k = 1,\ldots,8, \tag{3.4.36}$$

where $\varphi_k(\delta_1,\ldots,\delta_7)$ are power-series in the seven variables $\delta_k, k = 1, \ldots, 7$ starting with terms of degree ≥ 2. For a collision orbit we know **158**
that $q_4 = 0$, and hence $\delta_7 = 0$; and $p_k^* \to \bar{p}_k$, $q_k^* \to \bar{q}_k$, $k = 1,2,3$, as $t \to 0$, so that $\delta_k \to 0$, $k = 1,\ldots,6$, as $t \to 0$ or, equivalently, as $s \to \infty$. Moreover, the coefficients a_{kl} of the linear parts of the equations in the system (3.4.36), being functions of $\bar{p}_k, \bar{q}_k, k = 1,2,3$, alone, are functions determined uniquely by the three masses. Also $a_{k8} = 0$, $k = 1,\ldots 8$.

The nature of the solutions of the system (3.3.36) is related to the solution of the associated linear system

$$\delta_k' = \sum_{l=1}^{8} a_{kl}\delta_l$$

in view of the stability theory of solutions of ordinary differential equations and so we shall first study this simpler system. In order to study this system closely it is necessary to study the characterstic equation of the 8-rowed square matrix $A = (a_{kl})$ of the coefficients. Since $q_4^{*'} = \frac{1}{3} q_4^*$, we have $\delta_7' = \frac{1}{3}\delta_7$. Thus

$$a_{71} = 0,\ l \neq 7; a_{77} = \frac{1}{3}; a_{k8} = 0, k = 1, \ldots, 8.$$

So the matrix A of the coefficients of the linear part in (3.3.36) is

$$A = \begin{pmatrix} B & * & 0 \\ 0 & \frac{1}{3} & 0 \\ * & * & 0 \end{pmatrix},$$

where B is the 6-rowed square matrix of the coefficients $a_{kl}, k, l = 1, \ldots, 6$. If E_m denotes the m-rowed unit matrix, then the characterstic
159 polynomial of A is given by

$$|zE_8 - A| = z(z - \frac{1}{3})|zE_6 - B|. \tag{3.4.37}$$

In order to simplify the computation of the characterstic polynomial of A, we make a transformation $\delta_k = \sum_{l=1}^{8} c_{kl}\epsilon_l + \ldots$, with $|C| = \det(c_{kl}) \neq 0$, so that the system (3.3.36) is transformed into the system of differential equations (in matrix form):

$$\epsilon' = C^{-1}AC\epsilon + \chi(\epsilon_1, \ldots, \epsilon_8),$$

where χ_k are again power-series starting with quadratic terms. We wish to choose the transformation in such a way that the matrix $C^{-1}AC$ has a simple form. (Let us recall that the characteristic polynomial of A is the same as that of $C^{-1}AC$:

$$|zE - C^{-1}AC| = |C^{-1}(zE - A)C| = |zE - A|.$$

We shall now show that such a transformation can be obtained from the canonical transformation of the variables (ξ_k, η_k) to (p_k, q_k), in the following way. Let us recall that under this transformation, the variable p_4 does not appear in the new expression for E.

We consider the inverse of the canonical transformation from the variables (ξ_k, η_k) to the variables (p_k, q_k). This again is a canonical transformation and its Jacobian matrix is symplectic and has, in particular, a non-zero determinant. The transformation is given explicitly as follows:

$$\begin{aligned}&\xi_1 = p_1 c, \xi_2 = p_1 s, \xi_3 = p_2 c - p_3 s, \xi_4 = p_2 s + p_3 c;\\ &\eta_1 = q_1 c - q_o s, \eta_2 = q_1 s + q_o c,\ \eta_3 = q_2 c - q_3 s,\ \eta_4 = q_2 s + q_3 c,\end{aligned} \tag{3.4.38}$$

where $q_o = (q_4 - p_2 q_3 + p_3 q_2) p_1^{-1}$, $c = \cos p_4$, $s = \sin p_4$. **160**

Since the system of differential equations (3.4.35) was obtained by the substitution: $p_k = p_k^* t^{2/3}$, $q_k = q_k^* t^{-1/3}$, $k = 1, 2, 3$, $p_4 = p_4^*$, $q_4 = q_4^* t^{1/3}$, we have $q_o = q_o^* t^{-1/3}$. We introduce the variables ξ_k^*, η_k^* by setting

$$\xi_k = \xi_k^* t^{2/3}, \eta_k = \eta_k^* t^{-1/3}, k = 1, \dots, 4. \tag{3.4.39}$$

Under these substitutions we obtain from (3.3.38) the following transformation of the variables p_k^*, q_k^* to ξ_k^*, η_k^*:

$$\begin{aligned}&\xi_1^* = p_1^* c,\ \xi_2^* = p_1^* s, \xi_3^* = p_2^* c - p_3^* s, \xi_4^* = p_2^* s + p_3^* c,\\ &\eta_1^* = q_1^* c - q_o^* s,\ \eta_2^* = q_1^* s + q_o^* c,\ \eta_3^* = q_2^* c - q_3^* s, \eta_4^* = q_2^* s + q_3^* c.\end{aligned} \tag{3.4.40}$$

Then the substitutions $p_k^* = \overline{p}_k + \delta_k, q_k^* = \overline{q}_k + \delta_{k+3}$, $k = 1, 2, 3$, $p_4^* = \delta_8$, $q_4^* = \overline{q}_4 + \delta_7$, imply that ξ_k^*, η_k^* are functions of the variables $\delta_1, \dots, \delta_8$ and $\overline{\xi}_k, \overline{\eta}_k$ are obtained by setting $\delta_k = 0$, $k = 1, \dots, 8$. So we have

$$\begin{aligned}&\overline{\xi}_1 = \overline{p}_1,\ \overline{\xi}_2 = 0,\ \overline{\xi}_3 = \overline{p}_2,\ \overline{\xi}_4 = \overline{p}_3;\\ &\overline{\eta}_1 = \overline{q}_1,\ \overline{\eta}_2 = \overline{q}_o,\ \overline{\eta}_3 = \overline{q}_2,\ \overline{\eta}_4 = \overline{q}_3.\end{aligned} \tag{3.4.41}$$

These values have been determined explicitly in both the equilateral and **161**
the collinear cases. It is also clear that $\xi_k^* \to \overline{\xi}_k$, $\eta_k^* \to \overline{\eta}_k$, $k = 1, \dots, 4$, as $\delta_1 \to 0, \dots, \delta_8 \to 0$. Since to say that p_k^*, q_k^* satisfy the system of differential equations (3.4.35) is equivalent to saying that p_k, q_k satisfy

the Hamiltonian system, it follows that ξ_k, η_k satisfy the Hamiltonian system

$$\dot{\xi}_k = E_{\eta_k}, \dot{\eta}_k = -E_{\xi_k}, k = 1, \ldots, 4.$$

By means of the substitution (3.4.39) and $t = e^{-s}$, we obtain the following system of differential equations for ξ_k^*, η_k^*, considered as functions of the variable s:

$$\xi_k^{*'} = \frac{2}{3}\xi_k^* - E^*_{\eta_k^*}, \eta_k^{*'} = -\frac{1}{3}\eta_k^* + E^*_{\xi_k^*}, k = 1, \ldots, 4. \tag{3.4.42}$$

On differentiating E^* with respect to η_k^* we get

$$E^*\eta_k^* = \begin{cases} \left(\dfrac{1}{m_1} + \dfrac{1}{m_3}\right)\eta_k^* + \dfrac{1}{m_3}\eta_{k+2}^*, k = 1, 2; \\ \left(\dfrac{1}{m_2} + \dfrac{1}{m_3}\right)\eta_k^* + \dfrac{1}{m_3}\eta_{k-2}^*, k = 3, 4. \end{cases} \tag{3.4.43}$$

We take

$$\zeta_k^* = \xi_k^*, \ \zeta_{k+4}^* = -E^*_{\eta_k^*}, \ k = 1, \ldots, 4. \tag{3.4.44}$$

Then we can solve $\zeta_{k+4}^* = -E^*_{\eta_k^*}$, $k = 1, \ldots, 4$, and use (3.4.43) to express η_k^* in terms of ζ_{k+4}^*, $k = 1, \ldots 4$. We obtain

$$\eta_k^* = \begin{cases} -\dfrac{m_1}{m}(m_2 + m_3)\zeta_{k+4}^* + \dfrac{m_1 m_2}{m}\zeta_{k+6}^*, \ k = 1, 2; \\ \dfrac{m_1 m_2}{m}\zeta_{k+2}^* - \frac{m_2}{m}(m_1 + m_3)\zeta_{k+4}^*, \ k = 3, 4, \end{cases} \tag{3.4.45}$$

162 where $m = m_1 + m_2 + m_3$. The system (3.3.42) then becomes

$$\zeta_k^{*'} = \frac{2}{3}\zeta_k^* + \zeta_{k+4}^*, \zeta_{k+4}^{*'} = -\frac{1}{3}\zeta_{k+4}^* - F_k^*, \ k = 1, \ldots, 4, \tag{3.4.46}$$

where

$$F_k^* = \begin{cases} \dfrac{m_1 + m_3}{m_1 m_3} E^*_{\zeta_k^*} + \frac{1}{m_3} E^*_{\zeta_{k+2}^*}, \ k = 1, 2; \\ \dfrac{m_2 + m_3}{m_2 m_3} E^*_{\zeta_k^*} + \frac{1}{m_3} E^*_{\zeta_{k-2}^*}, \ k = 3, 4. \end{cases} \tag{3.4.47}$$

By the definition of ζ_k^* it is clear that the ζ_k^* tend to finite limits $\overline{\zeta}_k (k = 1, \dots, 8)$ and we have

$$\overline{\zeta}_1 = \overline{p}_1,\ \overline{\zeta}_2 = 0,\ \overline{\zeta}_3 = \overline{p}_2,\ \overline{\zeta}_4 = \overline{p}_3, \dots \tag{3.4.48}$$

where the $\overline{p}_k$ have already been determined. We now introduce the variables ϵ_k by setting

$$\zeta_k^* = \overline{\zeta}_k + \epsilon_k,\ k = 1, \dots, 8. \tag{3.4.49}$$

Then the system of differential equations (3.4.46) for ζ_k^* implies that the ϵ_k, considered as functions of the variable s, satisfy the system of differential equations

$$\epsilon_k' = \frac{2}{3}\epsilon_k + \epsilon_{k+4}, \epsilon_{k+4}' = -\frac{1}{3}\epsilon_{k+4} - \sum_{l=1}^{4} h_{kl}\epsilon_l + \dots, k = 1, \dots, 4, \tag{3.4.50}$$

where the coefficients h_{kl} are calcultated from (3.3.47) using the relations $E^*_{\zeta_k^*} = -U^*_{\xi_k^*}$, $k = 1, \dots, 4$, and substituting $\xi_k^* = \zeta_k^* = \overline{\zeta}_k + \epsilon_k$, $k = 1, \dots, 4$. Using the values of $\overline{\zeta}_k$ given by (3.4.48), we find that the matrix $H = (h_{kl})$, $k, l = 1, \dots, 4$, is given, in the equilateral case, by 163

$$H = r^{-3}\begin{pmatrix} \frac{m_2}{4} - 2(m_1 + m_3) & \frac{3\sqrt{3}}{4}m_2 & 0 & -\frac{3\sqrt{3}}{2}m_2 \\ \frac{3\sqrt{3}}{4}m_2 & m_1 + m_3 - \frac{5}{4}m_2 & -\frac{3\sqrt{3}}{2}m_2 & 0 \\ -\frac{9}{4}m_1 & -\frac{3\sqrt{3}}{4}m_1 & \frac{m_1 + m_2 + m_3}{4} & \frac{3\sqrt{3}}{4}(m_1 - m_2 - m_3) \\ -\frac{3\sqrt{3}}{4}m_1 & \frac{9}{4}m_1 & \frac{3\sqrt{3}}{4}(m_1 - m_2 - m_3) & -\frac{5}{4}(m_1 + m_2 + m_3) \end{pmatrix}$$

and in the collinear case, by

$$H = \rho^{-3}\begin{pmatrix} -2(m_1 + m_3) - 2m_2\omega^{-3} & 0 & 2m_2(\omega^{-3} - (1-\omega)^{-3}) & 0 \\ 0 & m_1 + m_3 + m_2\omega^{-3} & 0 & m_2((1-\omega)^{-3} - \omega^{-3}) \\ -2m_1(1 - \omega^{-3}) & 0 & -2m_1^{-3} - 2(m_2 + m_3)(1-\omega)^{-3} & 0 \\ 0 & m_1(1 - \omega^{-3}) & 0 & m_1\omega^{-3} + (m_2 + m_3)(1-\omega)^{-3} \end{pmatrix}$$

Then the matrix of the coefficients of the linear part of the system (3.4.50) is given by

$$\begin{pmatrix} \frac{2}{3}E_4 & E_4 \\ -H & -\frac{1}{3}E_4 \end{pmatrix},$$

and the characteristic polynomial of this matrix is **164**

$$\left|\left(z-\frac{2}{3}\right)\left(z+\frac{1}{3}\right)E_4+H\right|,$$

which is also the characteristic polynomial of the matrix A. Hence we have, recalling (3.4.37),

$$\left|\left(z-\frac{2}{3}\right)\left(z+\frac{1}{3}\right)E_4+H\right| = |zE_8-A| = \left(z-\frac{1}{3}\right)|zE_6-B|\,.$$

Denoting $\left(z-\frac{2}{3}\right)\left(z+\frac{1}{3}\right)$ by x, this gives

$$|xE_4+H| = \left(x+\frac{2}{9}\right)|zE_6-B|\,.$$

Explicit calculation of the left side shows that, in the equilateral case,

$$|xE_4+H| = \left(x+\frac{2}{9}\right)\left(x-\frac{4}{9}\right)\left(x^2-\frac{2}{9}x-\frac{8}{81}+\frac{1}{3}a\right),$$

where

$$a = (m_1m_2+m_2m_3+m_1m_3)\,(m_1+m_2+m_3)^{-2}, \tag{3.4.51}$$

and in the collinear case,

$$|xE_4+H| = \left(x+\frac{2}{9}\right)\left(x-\frac{4}{9}\right)\left(x+\frac{2}{9}+\frac{2}{9}b\right)\left(x-\frac{4}{9}-\frac{4}{9}b\right),$$

where

$$b = \frac{m_1(1+(1-\omega)^{-1}+(1-\omega)^2)+m_3(1+\omega^{-1}+\omega^{-2})}{m_1+m_2(\omega^{-2}+(1-\omega)^{-2})+m_3}. \tag{3.4.52}$$

Thus we obtain

$$|zE_6-B| = \left(x-\frac{4}{9}\right)\left(x^2-\frac{2}{9}x-\frac{8}{81}+\frac{1}{3}a\right),\ \text{in the equilateral case.}$$

$$\left(x - \frac{4}{9}\right)\left(x + \frac{2}{9} + \frac{2}{9}b\right)\left(x - \frac{4}{9} - \frac{4}{9}b\right), \text{ in the collinear case.}$$

165 It is clear from (3.4.51) that a is positive, and since $0 < \omega < 1$, b is also positive. The characteristic polynomial of B is a cubic in x and we shall determine all the roots. Consider first the equilateral case. Since $x = (z + \frac{1}{3})\,(z - \frac{2}{3})$, we have

$$x - \frac{4}{9} = z^2 - \frac{1}{3}z - \frac{2}{3} = (z-1)(z + \frac{2}{3}).$$

The roots of $x^2 - \frac{2}{9}x - \frac{8}{81} + \frac{1}{3}a = 0$ are $x = \frac{1}{9} \pm \frac{1}{3}\sqrt{1-3a}$, and hence we get $z^2 - \frac{1}{3}z - \frac{1}{3} \mp \frac{1}{3}\sqrt{1-3a} = 0$, of which the roots are $z = \frac{1}{6} \pm \frac{1}{6}\sqrt{13 \pm 12\sqrt{1-3a}}$. Here $0 \leq 1 - 3a < 1$ because we have $2(1-3a) = ((m_1-m_2)^2 + (m_2-m_3)^2 + (m_3-m_1)^2)(m_1+m_2+m_3)^{-2} \geq 0$, and $1 - 3a = 0$ if and only if $m_1 = m_2 = m_3$. So it follows that all the eigen values of the matrix B are real and are given by

$$\begin{aligned}
-a_o &= -\frac{2}{3}, -a_1 = +\frac{1}{6} - \frac{1}{6}\sqrt{13 + 12\sqrt{1-3a}},\\
-a_2 &= \frac{1}{6} - \frac{1}{6}\sqrt{13 - 12\sqrt{1-3a}},\\
a_3 &= \frac{1}{6} + \frac{1}{6}\sqrt{13 - 12\sqrt{1-3a}},\\
a_4 &= \frac{1}{6} + \frac{1}{6}\sqrt{13 + 12\sqrt{1-3a}}, a_5 = 1. \qquad (3.4.53)
\end{aligned}$$

Since $0 \leq 1 - 3a < 1$, we have $-a_2 \geq -a_1 > -a_o$, and all these are negative while $a_5 > a_4 \geq a_3 > 0$. All the six roots are distinct except when $1 - 3a = 0$, i.e. $m_1 = m_2 = m_3$, and then $a_1 = a_2$, $a_3 = a_4$.

Next we consider the collinear case. As before we get the 3 equations

$$x - \frac{4}{9} = (z-1)(z + \frac{2}{3}) = 0,\ \ x + \frac{2}{9} + \frac{2}{9}b = z^2 - \frac{1}{3}z + \frac{2}{9}b = 0,$$

$x-\frac{4}{9}-\frac{4}{9}b = z^2-\frac{1}{3}z-\frac{2}{3}-\frac{4}{9}b = 0$. Hence the eigen values of the matrix **166**
B are $z = 1$, $z = -\frac{2}{3}$, $z\frac{1}{6} \pm \frac{1}{6}\sqrt{1-8b}$, $z = \frac{1}{6} \pm \frac{1}{6}z = \sqrt{25+16b}$. Here $1-8b < 1$ and $1-8b$ can be negative, and in this case two of the eigen values are complex, complex conjugates of each other. For instance, if $m_1 = m_3$, then $\omega = \frac{1}{2}$ and $1-80b \gtreqless 0$ if $\frac{m_2}{m_1}\frac{55}{4}$. If $8b\ 1$, all the eigen values are real and distinct. If $8b = 1$, all the eigen values are real and there is a multiple root. Finally, then, the eigen values of the matrix B in the collinear case are given by:

$$
\begin{gathered}
-b_0 = -\frac{2}{3}, -b_1 = \frac{1}{6} - \frac{1}{6}\sqrt{25+16b},\ b_2 = \frac{1}{6} - \frac{1}{6}\sqrt{1-8b}, \\
b_3 = \frac{1}{6} + \frac{1}{+}\sqrt{1-8b},\ b_4 = \frac{1}{6} + \frac{1}{6}\sqrt{25+16b},\ b_5 = 1.
\end{gathered} \tag{3.4.54}
$$

It is clear that $-b_o$ and $-b_1$ are two negative roots and $-b_1 < -b_o$. There are four distinct positive roots if $8b < 1$, and four positive roots, two of them equal ($b_2 = b_3 = \frac{1}{6}$) if $8b = 1$, and two positive roots and a pair of complex conjugate roots with positive real parts if $8b > 1$.

In order to see how to utilize the knowledge of the eigenvalues for a study of the solutions of the system of equations (3.4.36), it is necessary to investigate in some detail the theory of stability of solutions of systems of ordinary differential equations.

5 Stability theory of solutions of differential equations

167 We shall now study the problem of the stability of the solutions of a system of ordinary differential equations of the first order. Let s be a real variable and $x_1, \ldots, x_m$ independent real variables. If f is a continuously differentiable function of s, we denote the derivative of f with respect to s by f'. We consider the system of m ordinary differential equations of the first order in m unknown real functions $x_k = x_k(s)$, $k = 1, \ldots, m$, of the variable s:

$$x'_k = \sum_{l=1}^{m} a_{kl} x_l + \varphi_k(x_1, \ldots, x_m), k = 1, \ldots, m, \tag{3.5.1}$$

where the φ_k are power-series in the m real variables $x_1, \ldots, x_m$ with real coefficients and starting with quadratic terms. The a_{kl} are real constants and we assume that the φ_k converge for $|x_k|$ sufficiently small. Let $\xi_1, \ldots \xi_m$ be given real numbers. We want to consider the problem of finding all solutions $x_k = x_k(s)$, $k = 1, \ldots, m$ of (3.5.1), taking the initial values $x_k(0) = \xi_k$ and studying their behaviour as $s \to \infty$. This is nearly the same as the equilibrium problem in mechanics, which is the following. Suppose that we have a mechanical system whose motion is governed by the system of equations (3.5.1). Since the right side of (3.5.1) contains no constant terms, $x_k(s) \equiv 0$ is a particular solution. The solution $x_k(s) \equiv 0$ is called an *equilibrium solution* of (3.5.1). If the right sides of (3.5.1) were power-series, possibly with constant terms, then $x_k(s) = c_k$, where $c_k, k = 1, \ldots, m$, are constants would be **168**
called an equilibrium solution if the c_k are a set of common zeros of the right sides. However, one sees easily that by taking the variables $X_k = x_k - c_k$ in place of x_k, one can reduce the system to one in the new variables, which is of the same form as (3.5.1), such that $X_k \equiv 0$ is an equilibrium solution for the new system. The solution $x_k = x_k(s)$, $k = 1, \ldots, m$, of (3.5.1) with initial values $x_k(0) = \xi_k$ defines a curve in m-dimensional Euclidean space, starting from the point $(\xi_1, \ldots, \xi_m)$. The problem of equilibrium consists in finding the behaviour of the solution when the initial values are varied in a sufficiently small neighbourhood of $(\xi_1, \ldots, \xi_m)$. Let s_o be a large positive number such that the solutions $x_k(s)$ of (3.5.1) have the property that $|x_k(s)|$, $k = 1, \ldots, m$, are sufficiently small for $s \geq s_o$, so that when these values are inserted in the power-series, the latter converge. Since the right side of (3.5.1) does not contain the variable s explicitly, the system remains unchanged if s is replaced by the variable $s - s_o$. We may assume that s has been replaced by $s - s_o$ and we then consider solutions in the half-line $s \geq 0$.

We have the following definition of stability of the solutions of (3.5.1). If for a given neighbourhood V of 0 in m-dimensional Euclidean

space we can find a neighbourhood W of 0 with $W \subset V$, such that for any point $(\xi_1, \ldots, \xi_m)$ in W, the solutions $x_k(s)$ of the system (3.5.1) taking initial values $x_k(0) = \xi_k, k = 1, \ldots, m$, exist for all $s \geq 0$ and remain in the neighbourhood V as $s \to \infty$, then the equilibrium solution of (3.5.1)
169 is called *stable*. If the equilibrium is not stable, we would also like to find for what initial conditions the solutions tend to zero as $s \to \infty$. The problem of stability of solutions of ordinary differential equations was first discussed by Poincaré by the method of power-series expansions, and independently by Liapounoff. But neither gave a method for obtaining all stable solutions. Bohl, and subsequently, Perron (Math. Zeit. (1928)) considered the problem of determining all stable solutions. Bohl studied the problem of stability also for systems of differential equations more general than the ones we consider, in the sense that the φ_k were assumed to be functions which satisfied certain growth conditions. Perron's method was simpler. We shall, however, give a treatment different from both these.

Consider the ball $\sum\limits_{k=1}^{m} x_k^2 < \epsilon$, where ϵ is a sufficiently small positive number. Instead of finding all solution of (3.5.1) which are asymptotic to 0 as $s \to \infty$, we consider the more general problem of finding all solutions $x_k(s)$ of (3.5.1) which, for all $s \geq 0$, belong to the ball $\sum\limits_{k=1}^{m} x_k^2 < \epsilon$. For this we start by simplifying the linear terms on the right side of (3.5.1) by a suitably chosen linear substitution. Let $A = (a_{kl})$ denote the m-rowed square matrix of the real coefficients $a_{kl}, k, l = 1, \ldots, m$, of the linear terms in (3.5.1) and we write (3.5.1) in the vector notation as

$$x' = Ax + \varphi(x). \tag{3.5.2}$$

We now apply the linear substitution

$$x = Cy \tag{3.5.3}$$

170 to x, where C is a real m-rowed square matrix with $|C| \neq 0$. Then the system (3.4.2) is transformed into the system

$$y' = C^{-1}ACy + \psi(y_1, \ldots, y_m), \tag{3.5.4}$$

where ψ is the column vector whose components ψ_k are power-series in the m independent variables $y_1, \ldots, y_m$ with real coefficients and starting with quadratic terms. Under the substitution (3.5.3), which is not necessarily orthogonal, the ball $\sum\limits_{k=1}^{m} x_k^2 < \epsilon$ is transformed into a bounded domain in m-dimensional y-space. Since $\epsilon > 0$ can be chosen sufficiently small, we may assume that this transformed domain is contained in the y-sphere $\sum\limits_{k=1}^{m} y_k^2 < \epsilon$. We choose the substitution (3.5.3) in such a way that the matrix $C^{-1}AC$ is in the normal form. (The reduction of a matrix to the normal form was first done by Weierstrass in 1868; it is what has subsequently been called the Jordan canonical form). Then the system of differential equations is reduced to a simpler form. For the moment we consider only the special case of (3.5.2) containing only linear terms on the right:

$$x' = Ax. \tag{3.5.5}$$

If all the eigenvalues $\lambda_1, \ldots, \lambda_m$ of the matrix A are real and distinct, then the matrix $C^{-1}AC$ is a diagonal matrix

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}$$

and (3.5.5) is reduced to the simple form **171**

$$y_k' = \lambda_k y_k, \ k = 1, \ldots, m. \tag{3.5.6}$$

This can be integrated immediatelyl to give the solution

$$y_k = c_k e^{\lambda_k s}, \ k = 1, \ldots, m, \tag{3.5.7}$$

where the c_k are constants of integration. Since the substitution (3.5.3) is linear, it is clear that if $x_k(s) \to 0 (k = 1, \ldots, m)$ as $s \to \infty$, then $y_k(s) \to 0 (k = 1, \ldots, m)$ too as $s \to \infty$, and conversely. Since we seek solutions which tend to zero as $s \to \infty$, we should have

$$c_k = 0 \text{ if } \lambda_k \geq 0 \text{ and } c_k \text{ arbitrary real if } \lambda_k < 0. \tag{3.5.8}$$

Hence we find that the general solution of (3.5.5) which goes to zero as $s \to \infty$ contains exactly the same number of arbitrary independent real parameters as the number of negative eigen-values of the matrix A. This motivates the conjecture that this result can be generalized to the system (3.5.2): suppose that all the eigen-values of A are real and distinct and exactly n of them are negative then there exist exactly n independent arbitrary real parameters in a general solution asymptotic to zero. This, however, is not in general true. It is true if zero is not an eigen-value of A. For the moment we shall consider only the case in which A has only real eigen-values. Later we shall generalize the result to the case in which some eigen-values are even complex.

172 We proceed to the following general theorem.

Theorem 3.5.1. *Suppose that all the eigen-values of the real matrix A area real, distinct and different from zero. If there are exactly n, $0 \leq n \leq m$, negative eigen-values, then a general solution of (3.5.2) which is such that $\sum^{m} x_k^2(s) < \epsilon$ for all $s \geq 0$ contains exactly n independent real parameters.*

Proof. Let $\lambda_1, \ldots, \lambda_m$ be the eigen-values of A and suppose that

$$\lambda_1 < 0, \ldots, \lambda_n < 0, \lambda_{n+1} > 0, \ldots, \lambda_m > 0. \tag{3.5.9}$$

It may happen that $n = 0$ or $n = m$. By a suitable choice of the substitution (3.5.3), we transform the system (3.5.2) into the system

$$y_k' = \lambda_k y_k + \psi_k(y_1, \ldots, y_m), k = 1, \ldots, m, \tag{3.5.10}$$

where the ψ_k are power-series with real coefficients and starting with quadratic terms in the variables $y_1, \ldots, y_m$ and convergent for small values of $|y_k|$. We may assume that ϵ is so small that the ψ_k converge in the y-sphere $\sum\limits_{k=1}^{m} y_k^2 < \epsilon$. In order to simplify the system (3.5.10) further, we introduce a non-linear substitution of the form

$$u_k = y_k - F_k(y_1, \ldots, y_n),\ k = 1, \ldots, m, \tag{3.5.11}$$

where the F_k are power-series with real coefficients in the n independent
173 variables $y_1, \ldots, y_n$ only, starting with quadratic terms and convergent

for small $|y_k|$. We remark that it is essential for our method that the F_k are functions of the variables $y_1, \ldots, y_n$ alone. If $n = 0$, then all $F_k \equiv 0$ and if $n = m$, then F_k are power-series in all the variables $y_1, \ldots, y_m$. It is immediately seen that the Jacobian matrix of the transformation (3.5.11) of the variables $y_1, \ldots, y_m$ to the variables $u_1, \ldots, u_m$ at the origin is the identity and hence the transformation is locally invertible at the origin. By local inversion we obtain y_k as power-series in the variables $u_1, \ldots, u_m$. In fact, if one considers the substitution (3.5.11) only for $k = 1, \ldots, n$, it defines a transformation of the variables $y_1, \ldots, y_n$ to $u_1, \ldots, u_n$. We can express $y_k, k = n+1, \ldots, m$, in terms of $u_1, \ldots, u_n$ by inserting in F_k the values of $y_1, \ldots, y_n$ in terms of $u_1, \ldots, u_n$ got by inversion from the first n equations in (3.5.11). Hence the inverse transformation of (3.5.11) has the same form. It is clear that if we make two such substitutions in succession, then the composite of the two is again such a substitution. So the substitutions of the form (3.5.11) form a group.

Differentiating (3.5.11) with respect to the variable s, we get

$$u'_k = y'_k - \sum_{l=1}^{n} F_{ky_l} y'_l.$$

Substituting for y'_k from (3.4.10) we obtain

$$u'_k = \lambda_k y_k + \psi_k(y_1, \ldots, y_m) - \sum_{l=1}^{n} F_{ky_l}(y_1, \ldots, y_n)(\lambda_l y_l + \psi_l(y)).$$

Once again it follows from (3.5.11) that $y_k = u_k + F_k, k = 1, \ldots, m$, so that we have, for $k = 1, \ldots, m$,

$$u'_k = \lambda_k u_k + \lambda_k F_k + \psi_k - \sum_{l=1}^{n} \lambda_l F_{ky_l} y_l - \sum_{l=1}^{m} F_{ky_l} \psi_l, \tag{3.5.12}$$

where the terms on the right can all be considered as functions of **174**
$u_1, \ldots, u_m$ after substituting for $y_1, \ldots, y_m$ the values obtained by the inverstion of (3.5.11). We set, for $k = 1, \ldots, m$,

$$\chi_k(u_1, \ldots, u_m) = \lambda_k F_k(y_1, \ldots, y_n) + \psi_k(y_1, \ldots, y_m) -$$

$$-\sum_{l=1}^{n}\lambda_l y_l F_{ky_l}(y_1,\ldots,y_n)-\sum_{l=1}^{n}\psi_l(y_1,\ldots,y_m)F_{ky_l}(y_1,\ldots,y_n), \tag{3.5.13}$$

so that we can write the system (3.5.12) in the form

$$u_k' = \lambda_k u_k + \chi_k(u_1,\ldots,u_m)\ k = 1,\ldots,m. \tag{3.5.14}$$

From the definitions (3.5.13) of the functions χ_k, it is clear that they are power-series with real coefficients, starting with quadratic terms in the variables $u_1,\ldots,u_m$, and convergent for small $|u_k|$. The idea of the proof now is to try to find the power-series $F_k(y_1,\ldots,y_n)$ in (3.5.11) in such a way that the power-series χ_k have simple forms. (If in (3.5.11) we took F_k as power-series in all the variables $y_1,\ldots,y_m$, then we could secure $\chi_k \equiv 0$, but it would be difficult to prove the convergence of F_k. However, with our choice of F_k, the χ_k may not all vanish identically, but the proof of convergence would be simpler). We take for F_k power-series in $y_1,\ldots,y_n$ with undetermined coefficients and try to find the coefficients in such a way that every term in χ_k contains at least one of the variables $u_{n+1},\ldots,u_m$ as a factor. That is, for $u_{n+1} = \ldots = u_m = 0$, we have

$$\chi_k(u_1,\ldots,u_n,\ 0,\ldots,0) \equiv 0, k = 1,\ldots,m. \tag{3.5.15}$$

175 In other words, we seek power-series $F_k(y_1,\ldots,y_n)$ so that the substitution (3.5.11) leads to the identity (3.5.15) for all $u_1,\ldots,u_n$ when we put $u_{n+1} = \ldots = u_m = 0$ in the power-series χ_k. We shall show that, under an additional condition, the power-series F_k are uniquely determined by the requirement (3.5.15). Then the system of differential equations (3.5.14) takes a simpler form in which it can be integrated directly. We shall later prove the convergence of the power-series F_k thus obtained.

Since $\epsilon > 0$ can be chosen as small as we want, the neighbourhood $\sum\limits_{k=1}^{m} y_k^2 < \epsilon$ of $y_1 = 0,\ldots,y_m = 0$ is transformed by the substitution (3.5.11) into a neighbourhood of $u_1 = 0,\ldots,u_m = 0$. Once again we may assume that this transformed neighbourhood is contained in the ball $\sum\limits_{k=1}^{m} u_k^2 < \epsilon$.

The conditions $u_{n+1} = 0, \ldots, u_m = 0$ mean that $y_{n+1}, \ldots, y_m$ satisfy the relations

$$y_k = F_k(y_1, \ldots, y_n),\ k = n+1, \ldots, m. \tag{3.5.16}$$

Substituting these power-series with undetermined coefficients in the expression (3.5.13) defining χ_k as a power-series in the variables $y_1, \ldots, y_n$, we get the functions

$$\chi_k(u_1, \ldots, u_n, 0, \ldots, 0)\ k = 1, \ldots, m,$$

where u_l denotes the power-series $u_l(y_1, \ldots, y_n, F_{n+1}(y_1, \ldots, y_n), \ldots, F_m(y_1, \ldots, y_n))$, $l = 1, \ldots, n$, in the variables $y_1, \ldots, y_n$. Since all the power-series F_k and χ_k start with quadratic terms, it follows that **176** $\chi_k(u_1, \ldots, u_n, 0, \ldots, 0)$ is a power-series starting with quadratic terms in the variables $y_1, \ldots, y_n$. We thus get power-series in $y_1, \ldots, y_n$ and these can be replaced, by local inversion of the substitution (3.5.11), by series in $u_1, \ldots, u_m$. It is however not necessary to use the inversion of (3.5.11), and so not necessary to use the variables u_k at all. Instead, one can direectly consider the condition (3.5.15) to hold identically as power-series in $y_1, \ldots, y_n$. This implies certain polynomial relations for the coefficients of F_k. The coefficients can then be determined by induction from these relations, in the following way.

Let g be an integer ≥ 2. Suppose that the coefficients of all terms of total degrees $2, 3, \ldots, g-1$ in $F_k (k = 1, \ldots, m)$ have been determined. Then we show that the coefficients of the terms of total degree g can be determined. Consider a term of total degree g in F_k, of the form

$$c_k y_1^{g_1} \ldots y_n^{g_n}, \tag{3.5.17}$$

where $g_1, \ldots, g_n$ are non-negative integers such that $g_1 + \ldots + g_n = g \geq 2$. The (real) coefficients c_k in (3.5.17) is determined by equating to zero the coefficient of the term $y_1^{g_1} \ldots y_n^{g_n}$ in the power-series χ_k obtained on replacing $y_l, l = n+1, \ldots, m$, by the power-series F_l. Then $c_k g_l y_1^{g_1} \ldots y_1^{g_l - 1} \ldots y_n^{g_n}$ is a term of total degree $g-1$ in F_{ky_l}, so that the coefficient of $y_1^{g_1} \ldots y_n^{g_n}$ in $y_l F_{ky_l}$ is $c_k g_l$. Then in view of (3.5.13) the **177** condition (3.5.15) implies that

$$c_k(\lambda_k - \sum_{l=1}^{n} g_l\lambda_l) = \text{ coefficient of } y_1^{g_1}\ldots y_n^{g_n} \text{ in } (-\psi_k + \sum_{l=1}^{n} F_{ky_l}\psi_l). \tag{3.5.18}$$

Since ψ_k starts with quadratic terms in $y_1, \ldots, y_n$ and

$$\psi_k(y_1, \ldots, y_m) = \psi_k(y_1, \ldots, y_n, F_{n+1}, \ldots, F_m),\ k = 1, \ldots, m,$$

where again $F_l, l = n+1, \ldots, m$, are power-series starting with quadratic terms in $y_1, \ldots, y_n$, the coefficient of $y_1^{g_1}\ldots y_n^{g_n}$ in ψ_k involves only the coefficients in F_k of total degrees $2, 3, \ldots, g-1$, which are known by the induction assumption, and is actually a polynomial in these known coefficients of total degrees $\leq g-1$. The same is the case with the coefficient of $y_1^{g_1}\ldots y_n^{g_n}$ in $\sum_{l=1}^{n} F_{ky_l}\psi_l$. Thus the right side of (3.5.18) is known by the induction assumption.

In the particular case $g = 2$, we have only to consider the contribution from the quadratic terms in $-\psi_k$, because $F_{ky_l}\psi_l$ is a power-series starting with cubic terms in the variables $y_1, \ldots, y_n$. Moreover, the quadratic terms in $-\psi_k$ in the variables $y_1, \ldots, y_n$ give contributions only from the terms $y_py_q,\ p, q = 1, \ldots, n$, and not from the terms involving $y_pF_q,\ p, q = n+1, \ldots, m$. Hence in this case the right side of (3.5.18) is completely determined by the coefficients of ψ_k itself, and hence we can start the induction.

178 The coefficients c_k can then be determined from (3.5.18) whenever

$$\lambda_k \neq \sum_{l=1}^{n} g_l\lambda_1,\ k = 1, \ldots, m, \tag{3.5.19}$$

where g_l are non-nagative integers with $g_1 + \ldots + g_n = g \geq 2$. We see that (3.5.19) is actually only a finite set of conditions to be satisfied by $\lambda_1, \ldots, \lambda_m$. In fact, $-\lambda_l > 0$, $1 = 1, \ldots, n$, implies that $-\sum_{l=1}^{n} g_l\lambda_l \to \infty$ as $g \to \infty$. This means that for integers $g_1, \ldots, g_n \geq 0$ such that $g_1 + \ldots + g_n = g \geq g_o$, where g_o is sufficiently large, the condition (3.5.19) is automatically satisfied, and so we need assume (3.5.19) only for a finite set of n-tuples of integers $g_1, \dots, g_n$, all non-negative, with $2 \leq g < g_o$, $k = 1, \dots, n$. On the other hand, if $k = n+1, \dots, m$, we know that $\lambda_k > 0$ and $-\sum\limits_{l=1}^{n} g_l \lambda_l > 0$, so that (3.5.19) is always satisfied. Then it follows that all the coefficients in the power-series F_k can be determined by induction, provided that the finite set of conditions (3.5.19) holds. (If the condition (3.5.19) does not hold, there are complications which, however, can be overcome, as we shall show later).

We remark that if we had taken F_k to be power-series in all the variables $y_1, \dots, y_m$, then $\lambda_k - \sum\limits_{l=1}^{m} g_l \lambda_l$ would become arbitrarily small if $0 < n < m$, and so we would have arbitrarily small denominators for the coefficients c_k to be determined from (3.5.18), which would make the proof of convergence more difficult.

We shall prove the convergence of the power-series F_k obtained above later by the Cauchy method of majorants. We proceed with the proof of the theorem assuming for the moment the convergence of the power-series F_k for sufficiently small $|y_l|$, $l = 1, \dots, n$. **179**

We shall now determine the general solution of the system of differential equations (3.5.14):

$$u'_k = \lambda_k u_k + \chi_k, \ k = 1, \dots, m,$$

where the power-series χ_k are determined by (3.5.13) after inserting the power-series $F_k = F_k(y_1(u), \dots, y_n(n))$, $k = n+1, \dots, m$, obtained above; we seek only solutions u_k such that $\sum\limits_{k=1}^{m} u_k^2 < \epsilon$ for all $s \geq 0$. First we show that $u_k \equiv 0$ for $k = n+1, \dots, m$. For this we set

$$v = \sum_{k=n+1}^{m} u_k^2. \tag{3.5.20}$$

If $n = m$, this sum is empty and there is nothing to prove. So let $n < m$. Now, if $u_k = u_k(s)$ are solutions of the system $u'_k = \lambda_k u_k + \chi_k$, $k = n+1, \dots, m$, then we obtain for v the differential equation

$$v' = 2 \sum_{k=n+1}^{m} u_k u'_k = 2 \sum_{k=n+1}^{m} \lambda_k u_k^2 + 2 \sum_{k=n+1}^{m} u_k \chi_k.$$

We estimate the right side from below. Let $\lambda = \min(\lambda_{n+1}, \ldots, \lambda_m)$. Then $\lambda > 0$ and $2 \sum_{k=n+1}^{m} \lambda_k u_k^2 \geq 2\lambda v$. Since $\chi_k(u_1, \ldots, u_n,\ 0, \ldots, 0) \equiv 0$ iden-
180 tically in $u_1, \ldots, u_n$, and χ starts with quadratic terms, at least one of $u_{n+1}, \ldots, u_m$ occurs as a factor in each term of $\chi_k(u_1, \ldots, u_m)$. Since u_k is real, $u_k^2 \leq v$, $k = n+1, \ldots, m$, and so $|u_k| \leq \sqrt{v}$. Consequently, each term of $u_k \chi_k$, $k = n+1, \ldots, m$, has a factor of the form u_p, u_q, $p, q = n+1, \ldots, m$, and hence of absolute value $\leq v$ whereas the remaining factor of the term is a product of powers of $u_1, \ldots, u_m$, at least of total degree 1. As the χ_k are uniformly convergent, one can choose $\epsilon > 0$ so small that $\sum_{k=1}^{m} u_k^2 < \epsilon$ implies $|2 \sum_{k=n+1}^{m} u_k \chi_k| \leq \lambda v$. Hence, in particular, we have $2 \sum_{k=n+1}^{m} u_k \chi_k \geq -\lambda v$, so that we obtain the differential inequality $v' \geq \lambda v$ and hence $(ve^{-\lambda s})' \geq 0$. So $ve^{-\lambda s}$ is a non-decreasing function of s in $s \geq 0$. As $s \to \infty$, $v(s)$ remains bounded since $v(s) \leq \sum_{k=1}^{m} u_k^2 < \epsilon$. But $\lambda > 0$, so $e^{-\lambda s}$ and hence $ve^{-\lambda s} \to 0$ as $s \to \infty$. Since $ve^{-\lambda s}$ is non-negative and non-decreasing, we should have $ve^{-\lambda s} \equiv 0$ and hence $v \equiv 0$. This means that $u_k(s) = 0$, $k = n+1, \ldots, m$, which proves our assertion.

Now in view of (3.5.13), then system (3.5.14) reduces to $u_k' = \lambda_k u_k$, $k = 1, \ldots, n$. On integration we obtain a general solution of (3.5.14) with $\sum_{k=1}^{m} u_k^2 < \epsilon$ and this is given by

$$u_k = c_k e^{\lambda_k s},\ k = 1, \ldots, m;\ c_k = 0,\ k = n+1, \ldots, m.$$

Since $c_k = u_k(0)$, it follows that $\sum_{k=1}^{n} c_k^2 < \epsilon$. Conversely, given the initial conditions $u_k(0) = c_k$ with $\sum_{k=1}^{n} c_k^2 < \epsilon$, since $\lambda_k < 0$ for $k = 1, \ldots, n$,
181 any general solution of (3.5.14) necessarily satisfies $\sum_{k=1}^{m} u_k^2 < \epsilon$. On the other hand, by the uniqueness of the solutions of systems of differential equations with prescribed initial conditions, we see that, given $c_k = u_k(0)$ with $\sum_{k=1}^{n} c_k^2 < \epsilon$, we have determined the unique solution

of (3.5.14) with these initial values asymptotic to zero. Thus we have determined all solutions of (3.5.14) with the property that $\sum_{k=1}^{m} u_k^2 < \epsilon$ for all $s \geq 0$. Going back to the differential equations satisfied by the unknown functions $y_k(s)$ by means of the inverse of the substitution (3.5.13) we can express y_k in the form $y_k = u_k + G_k(u_1, \ldots, u_n)$, $k = 1, \ldots, m$, where the G_k are power-series in $u_1, \ldots, u_m$ without linear terms. In fact, let b_k denote the initial values of y_k, $k = 1, \ldots, m$. Then $u_k(0) = 0$, $k = n + 1, \ldots, m$, is equivalent to $b_k = F_k(b_1, \ldots, b_n)$, $k = n + 1, \ldots, m$. We can choose $b_1, \ldots, b_n$ arbitrarily with the only condition that if $c_k = b_k - F_k(b_1, \ldots, b_n)$, $k = 1, \ldots, n$, then $\sum_{k=1}^{n} c_k^2 < \epsilon$. Thus the initial values $b_1, \ldots, b_m$ have to satisfy $m - n$ conditions $b_k = F_k(b_1, \ldots, b_n)$, $k = 1, \ldots, m$. If we prove that the F_k are convergent power-series, then it follows that the initial values for $y_k(s)$ satisfy $m - n$ analytical relations. So the solutions $y_k(s)$ with $\sum_{k=1}^{m} y_k^2 < \epsilon$ for all $s \geq 0$ lie on an $(m-n)$-dimensional analytic manifold defined by the equations $y_k = F_k(y_1, \ldots, y_n)$, $k = n + 1, \ldots, m$, and we have a (local) parametric representation for this manifold. The solutions are given explicitly by

$$y_k = c_k e^{\lambda_k s} + G_k(c_1 e^{\lambda_1 s}, \ldots, c_n e^{\lambda_n s}),\ k = 1, \ldots, n;$$
$$y_k = G_k(c_1 e^{\lambda_1 s}, \ldots, c_n e^{\lambda_n s}),\ k = n + 1, \ldots, m.$$

Finally we go back to the variables x_k by means of the inverse of the **182**
linear substitution $x = C_y$, $|C| \neq 0$. We see therefore that if $a_1, \ldots, a_m$ denote the initial values $x_1(0), \ldots, x_m(0)$ of the solution asymptotic, as $s \to \infty$, to the equilibrium solution of the system (3.5.1), then a_k also satisfy $m - n$ analytic relations. Thus the solutions $x_k(s)$, $k = 1, \ldots, m$, asymototic to the equilibrium solution fill an $(m - n)$-dimensional analytic manifold in m-dimensional Euclidean space. The general solution then involves n real parameters and we have proved the theorem, but for the convergence of the power-series, subject to the condition (3.5.18).

Now we shall proceed to prove the convergence of the power-series F_k by Cauchy's method of majorants. We have used this method earlier to prove Theorem 1.3.1. It is a little more difficult in the present case, since the equations (3.5.13) defining χ_k involve the partial derivatives

F_{ky_l} and hence the condition (3.5.15) will give a system of partial differential equations for F_k, $k = 1, \ldots, m$. For determining the power-series F_k, we use the condition (3.4.15) leading to

$$\lambda_k F_k - \sum_{l=1}^{n} \lambda_l y_l F_{ky_l} = -\psi_k(y_1, \ldots, y_n,\ F_{n+1}, \ldots, F_m) + \\ + \sum_{l=1}^{n} F_{ky_l}(y_1, \ldots, y_n)\psi_l(y_1, \ldots, y_n, F_{n+1}, \ldots, F_m); \qquad (3.5.21)$$

183 which implies recurrence relations for the coefficients in $F_k, k = 1, \ldots, m$. Comparing the coefficients of a typical term $c_k y_1^{g_1} \ldots y_n^{g_n}$ of total degree $g = g_1 + \ldots + g_n \geq 2$, we obtained the relation (3.5.18), which may be re-written as

$$c_k(\lambda_k - \sum_{l=1}^{n} g_l \lambda_l) = \left\{ -\psi_k + \sum_{l=1}^{n} \psi_l F_{ky_l} \right\}_{g_1 \ldots g_n} \quad (k = 1, \ldots, m), \qquad (3.5.22)$$

where $\{f\}_{g_1,\ldots,g_n}$ stands for the coefficient of the term $y_1^{g_1} \ldots y_n^{g_n}$ in the power-series f in $y_1, \ldots, y_n$. Under the assumption (3.5.19), the coefficients in F_k are determined recursively by (3.5.22). In order to obtain majorants for F_k, we estimate c_k in the following way. From (3.5.22) we have

$$|c_k||\lambda_k - \sum_{l=1}^{n} g_l \lambda_l| = \left| \left\{ -\psi_k + \sum_{l=1}^{n} \psi_1 F_{ky_l} \right\}_{g_1 \ldots g_n} \right|, k = 1, \ldots m. \qquad (3.5.23)$$

Let $\alpha = \min(-\lambda_1, \ldots, -\lambda_n)$. Then $\alpha > 0$ and we can write

$$\lambda_k - \sum_{l=1}^{n} \lambda_l g_l \geq \lambda_k + \alpha(g_1 + \ldots g_n) = \lambda_k + \frac{\alpha}{2}(g_1 + \ldots + g_n) + \frac{\alpha}{2}(g_1 + \ldots + g_n).$$

For sufficiently large $g = g_1 + \ldots + g_n$, we have $\lambda_k + \dfrac{\alpha}{2} g > 0$ and hence for such $g_1, \ldots, g_n$, $\lambda_k - \sum\limits_{l=1}^{n} g_l \lambda_l > \dfrac{\alpha}{2}(g_1 + \ldots + g_n)$. Since by (3.5.19), $\lambda_k - \sum\limits_{l=1}^{n} g_l \lambda_l \neq 0$ and $\lambda_k + \dfrac{1}{2}\alpha g \leq 0$ for only finitely many n-tuples

$(g_1, \dots, g_n)$, we can find a sufficiently large positive constant γ_1 such that always

$$|\lambda_k - \sum_{l=1}^{n} g_l \lambda_l| > \gamma_1^{-1} g, \; k = 1, \dots, m.$$

We shall hereafter denote by $\gamma_2, \gamma_3, \dots$ sufficiently large positive constants. It follows now from (3.5.23) that **184**

$$|c_k|(g_1 + \dots + g_n) \leq |\gamma_1 \left\{ -\psi_k + \sum_{l=1}^{n} \psi_l F_{ky_l} \right\}_{g_1 \dots g_n} |, \; k = 1, \dots, m. \tag{3.5.24}$$

We know that ψ_k are power-series starting with quadratic terms and converging in a complex neighbourhood of $y_1 = 0, \dots, y_m = 0$, say, $|y_1| \leq \rho_1, \dots, |y_m| \leq \rho_m$. Suppose that $|\psi_k| \leq \gamma_2$ in this region, $k = 1, \dots, m$. If $h_1, \dots, h_m$ are non-negative integers, then by Cauchy's formula

$$\left| \{\psi_k\}_{h_1 \dots h_m} \right| \leq \gamma_2 \rho_1^{-h_1} \dots \rho_m^{-h_m}.$$

So we can write

$$\psi_k \prec \sum_{h_1 + \dots + h_m \geq 2} \gamma_2 \frac{y_1^{h_1} \dots y_m^{h_m}}{\rho_1^{h_1} \dots \rho_m^{h_m}}.$$

If $h = h_1 + \dots h_m$, then

$$\psi_k \prec \sum_{h=2}^{\infty} \gamma_2 \sum_{h_1 + \dots + h_m = h} \left(\frac{y_1}{\rho_1} \right)^{h_1} \dots \left(\frac{y_m}{\rho_m} \right)^{h_m} \prec \sum_{h=2}^{\infty} \gamma_2 \left(\frac{y_1}{\rho_1} + \dots \frac{y_m}{\rho_m} \right)^{h}. \tag{3.5.25}$$

Let $\gamma_3^{-1} = \min(\rho_1, \dots, \rho_m)$. Then $\gamma_3 > 0$ and

$$\frac{y_1}{\rho_1} + \dots + \frac{y_m}{\rho_m} \prec \gamma_3 (y_1 + \dots + y_m),$$

so that the right side of (3.5.25) is majorized by $\gamma_2 \sum\limits_{h=2}^{\infty} \gamma_3^h (y_1 + \dots + y_m)^h$, which is the formal power-series **185**

$$\gamma_2 \frac{\gamma_3^2 (y_1 + \dots + y_m)^2}{1 - \gamma_3 (y_1 + \dots y_m)} \equiv \Psi(y_1, \dots, y_m), \text{ say.}$$

Then $\psi_k \prec \Psi$, $k = 1, \ldots, m$. Let F_k^* denote the power-series obtained by replacing the coefficients in F_k by their absolute values. Now substituting Ψ for each ψ_k and F_k^* for F_k we obtain a majorant for the power-series on the right side of (3.5.21). In fact, since the coefficients in F_k^* are non-negative, the coefficients in $F_{ky_l}^*$ are also non-negative and hence the right side of (3.5.24) is majorized by

$$\left\{\gamma_4(1 + \sum_{l=1}^{n} F_{ky_l}^*(y_1, \ldots, y_n))\Psi(y_1, \ldots, y_n, F_{n+1}^*, \ldots, F_m^*)\right\}_{g_1 \cdots g_n},$$

which implies that

$$|c_k|(g_1 + \ldots g_n) \leq \left\{\gamma_4\left(1 + \sum_{l=1}^{n} F_{ky_l}^*\right)\Psi\left(y_1, \ldots, y_n, F_{n+1}^*, \ldots, F_m^*\right)\right\}_{g_1 \cdots g_n}.$$

Since $|c_k|$ is the coefficient of $y_1^{g_1} \ldots y_n^{g_n}$ in F_k^*, we see that $|c_k| g_l$ is the coefficient of $y_1^{g_1} \ldots y_l^{g_1 - 1} \ldots y_n^{g_n}$ in $F_{ky_l}^*$, and hence of $y_1^{g_1} \ldots y_n^{g_n}$ in $y_l F_{ky_l}^*$. In other words,

$$\left\{\sum_{l=1}^{n} y_l F_{ky_l}^*\right\}_{g_1 \cdots g_n} = |c_k|(g_1 + \ldots + g_n),$$

and thus we get the majorization

$$\sum_{l=1}^{n} y_l F_{ky_l}^* \prec \gamma_4\left(1 + \sum_{l=1}^{n} F_{ky_l}^*\right)\Psi(y_1, \ldots, y_n, F_{n+1}^*, \ldots, F_m^*). \tag{3.5.26}$$

186 Let $G_1, \ldots, G_m$ be m power-series with non-negative undetermined coefficients, starting with quadratic terms, in the variables $y_1, \ldots, y_n$, satisfying the m partial differential equations

$$\sum_{l=1}^{n} y_l G_{ky_l} = \gamma_4\left(1 + \sum_{l=1}^{n} G_{ky_l}\right)\Psi(y_1, \ldots, y_n,\ G_{n+1}, \ldots, G_m),$$
$$k = 1, \ldots, m. \tag{3.5.27}$$

Let $d_k = \{G_k\}_{g_1 \ldots g_n}$. Then comparing the coefficients of $y_1^{g_1} \ldots y_n^{g_n}$ on both sides of (3.5.27), we have

$$d_k(g_1 + \ldots g_n) = \gamma_4 \left\{(1 + \sum_{l=1}^{n} G_{ky_l})\Psi(y_1, \ldots, y_n,\ G_{n+1}, \ldots, G_m)\right\}_{g_1 \ldots g_n} \tag{3.5.28}$$

As before, the relation (3.5.28) is a recurrence relation for determining the coefficients d_k of G_k. For $g = 2$, the right side of (3.5.28) contains only coefficients of quadratic terms of the form $y_p y_q,\ p, q = 1, \ldots, n$, alone in Ψ, since Ψ starts with quadratic terms so that all the terms of $G_{ky_l}\Psi, y_p G_q$, $p = 1, \ldots, n$; $q = n+1, \ldots, m$, or $G_p G_q$, $p, q = n+1, \ldots, m$, are of total degree ≥ 2. Hence the right side of (3.5.28) is known in this case, and therefore also the coefficients of the quadratic terms in G_k. For $g > 2$, we see, as in the determination of the coefficients of F_k, that the right side of (3.5.28) involves coefficients in G_k of terms of total degrees $2, \ldots, g - 1$, which are already determined, and thus all the coefficients in G_k are uniquely determined by induction. So the power-series G_k are uniquely determined by (3.5.27).

We next prove by induction that $F_k^* \prec G_k$, $k = 1, \ldots, m$. If $g =$ 187
$g_1 + \ldots + g_n = 2$, then we see immediately from (3.5.26) that

$$\begin{aligned} 2\left\{F_k^*\right\}_{g_1 \ldots g_n} &\leq \left\{\gamma_4 \left(1 + \sum_{l=1}^{n} F_{ky_l}^*\right)\Psi(y_1, \ldots, y_n,\ F_{n+1}^*, \ldots, F_m^*)\right\}_{g_1 \ldots g_n} \\ &= \left\{\gamma_4 \Psi(y_1, \ldots, y_n, F_{n+1}^*, \ldots, F_m^*)\right\}_{g_1 \ldots g_n}. \end{aligned}$$

But by the above construction, the right side is precisely the coefficient $2\{G_k\}_{g_1 \ldots g_n}$. Suppose that for all non-negative integers $g_1, \ldots, g_n$ with $2 \leq g_1 + \ldots + g_n \leq g - 1$ we have

$$\{F_k^*\}_{g_1 \ldots g_n} \leq \{G_k^*\}_{g_1 \ldots g_n},\ (k = 1, \ldots, m). \tag{3.5.29}$$

Then we shall show that for all non-negative integers $h_1, \ldots, h_n$ with $h_1 + \ldots + h_n = g$, we have

$$\{F_k^*\}_{h_1 \ldots h_n} \leq \{G_k^*\}_{h_1 \ldots h_n}.$$

By (3.5.36) we have

$$\{F_k^*\}_{h_1\ldots h_n}(h_1+\ldots+h_n) \leq \left\{\gamma_4\left(1+\sum_{l=1}^{n} F^*_{ky_l}\right)\Psi(y_1,\ldots,y_n,F^*_{n+1},\ldots,F^*_m)\right\}_{h_1\ldots h_n}. \quad (3.5.30)$$

On the right side only coefficients of terms of total degree $\leq g-1$ occur, and for these coefficients, (3.5.29) holds. In other words, the right side of (3.5.50) is majorized by the coefficient

$$\left\{\gamma_4\left(1+\sum_{l=1}^{n} G_{ky_l}\right)\Psi(y_1,\ldots,y_n,\ G_{n+1},\ldots,G_m)\right\}_{h_1\ldots h_n}.$$

188 But by the construction of the power-series G_k, this is equal to $\{G_k\}_{h_1\ldots h_n}$ $(h_1+\ldots+h_n)$ and this, by induction, proves our assertion. So in order to prove the convergence of F_k^*, and hence of F_k, it is enough to prove the convergence of G_k.

It is easy to see that $G_1 = \ldots = G_n$. In fact, consider the power-series G with non-negative undetermined coefficients starting with quadratic terms, satisfying the partial differential equation

$$\sum_{l=1}^{n} y_l G_{y_l} = \gamma_4\left(1+\sum_{l=1}^{n} G_{y_l}\right)\Psi(y_1,\ldots,y_n,G,\ldots,G). \quad (3.5.31)$$

The coefficients of G can be determined by induction as in the case of G_k. We have already remarked that the coefficients of the quadratic terms are the same in all the G_k and they are obtained by the contributions from the quadratic terms of the type y_py_q, $p,q = 1,\ldots,n$, in Ψ alone, and that there is no contribution either from terms of the type $y_pG_q, p = 1,\ldots,n; q = n+1,\ldots,m$, or from terms of the type G_pG_q, $p,q = n+1,\ldots,m$. But these are exactly the coefficients of the corresponding quadratic terms of G. Then using the recurrence formula (3.5.28), we see by induction that all the corresponding coefficients of G_k are equal, and equal to those of G. Hence, $G = G_1 = \ldots = G_m$ is uniquely determined by (3.5.31).

If we set $y_1 = \ldots = y_n = y$ in the power-series $G(y_1, \ldots, y_n)$, we obtain a power-series with non-negative coefficients in one variable y, starting with quadratic term; we shall denote this by $H(y)$. If $H(y)$ converges for some positive value of y, then it is clear that $G(y_1, \ldots, y_n)$ **189**
converges for $|y_1| < y, \ldots, |y_n| < y$. In fact,

$$\left|\{G(y_1, \ldots, y_n)\}_{g_1 \ldots g_n}\right| \leq \{H(y)\}_{g_1 + \ldots + g_n}.$$

Since $H(y) = G(y, \ldots, y)$, it follows from (3.5.31) and the definition of Ψ that we have

$$\begin{aligned} yH_y &= \gamma_4(1 + H_y)\Psi(y, \ldots, y, H, \ldots, H) \\ &= \gamma_5(1 + H_y)\frac{(ny + (m-n)H)^2}{1 - \gamma_3(ny + (m-n)H)}. \end{aligned}$$

The right side of this can be majorized further as follows. Since

$$\frac{(ny + (m-n)H)^2}{1 - \gamma_3(ny + (m-n)H)} = \gamma_3^{-2}\sum_{l=2}^{\infty}\gamma_3^l(ny + (m-n)H)^l,$$

we have

$$\frac{(ny + (m-n)H)^2}{1 - \gamma_3(ny + (m-n)H)} \prec \gamma_3^{-2}\sum_{l=2}^{\infty}(\gamma_3 m)^l(y + H)^l = \gamma_3^{-2}\frac{(\gamma_3 m)^2(y + H)^2}{1 - \gamma_3 m(y + H)}.$$

Putting $\gamma_3 m = \gamma_6$ and $\dfrac{(y + H)^2}{1 - \gamma_6(y + H)} = \Phi(y, H)$, we get the majorization

$$yH_y \prec \gamma_7(1 + H_y)\Phi(y, H).$$

Let $J = J(y)$ be the power-series in y starting with the second degree term, satisfying the differential equation

$$y\, J_y = \gamma_7(1 + J_y)\Phi(y, J). \tag{3.5.32}$$

The coefficients in the power-series J can be determined by induction **190**
on comparing coefficients on both sides. It is easy to see, as in the case of G and H, that $J \succ H$, and hence it is enough to prove the convergence

of J. One can integrate (3.5.32) and obtain J directly. However, one can majorize J by the following simple method. Let

$$J(y) = \sum_{k=2}^{\infty} a_k y^k; \tag{3.5.33}$$

J starts with quadratic terms since $\dfrac{\gamma_3^2(y_1 + \ldots + y_m)^2}{1 - \gamma_3(y_1 + \ldots + y_m)}$ does and hence also H does. Then on comparing the coefficients of y^k on both sides of (3.5.32), we obtain

$$ka_k = \left\{\gamma_7\left(1 + \sum_{l=2}^{\infty} la_l y^{l-1}\right)\Phi(y, J)\right\}_k \quad (k = 2, 3, \ldots). \tag{3.5.34}$$

The right side of (3.5.34) has contributions only from coefficients a_l of terms of degree at most $k - 1$. In other words, it involves only the coefficients $a_2, \ldots, a_{k-1}$ and so $0 < \dfrac{l}{k} < 1$. From (3.5.34),

$$a_k = \left\{\gamma_7\left(\frac{1}{k} + \sum_{l=2}^{k-1} \frac{l}{k} a_l y^{l-1}\right)\Phi(y, J)\right\}_k.$$

If we take

$$a_k^* = \left\{\gamma_7\left(1 + \sum_{l=2}^{k-1} a_l^* y^{l-1}\right)\Phi(y, J)\right\}_k, \ k = 2, 3, \ldots, \tag{3.5.35}$$

then again by induction $a_k \leq a_k^*$ for all $k \geq 2$. We get a power-series
191 in one variable y with non-negative coefficients and starting with the second degree term by setting

$$K(y) = \sum_{k=2}^{\infty} a_k^* \, y_k.$$

Then $J \prec K$. We see that the relations (3.5.35) defining a_k^* inductively imply that

$$K = \gamma_7(1 + y^{-1}K)\frac{(y + K)^2}{1 - \gamma_6(y + K)}.$$

Since K starts with quadratic terms, $K_1 = y^{-1}K$ is a power-series starting with linear terms and moreover satisfies the identity

$$K_1 = \frac{\gamma_7 y(1+K_1)^3}{1-\gamma_6 y(1+K_1)}. \tag{3.5.36}$$

This is again obtained in a constructive way, and it is now enough to prove the convergence of K_1, which implies the convergence of K, and so that of J and hence of G_k, F_k also.

In order to obtain a solution of the algebraic equation (3.5.36) for K_1 in a convergent power-series, one may use the implicit function theorem. (It is easily verified that the conditions of this theorem are satisfied). We shall, however, prove the convergence of K_1 directly without determining a majorant of K_1 explicitly. For this purpose we construct a simpler power-series which majorizes K_1. We can write

$$\begin{aligned} K_1 &= \gamma_7 y(1+K_1)^3 \sum_{l=30}^{\infty} (\gamma_6 y)^l (l+K_1)^l = \gamma_7 y \sum_{l=0}^{\infty} (\gamma_6 y)^l (1+K_1)^{l+3} \\ &= \gamma_7 y \sum_{l=0}^{\infty} (\gamma_6 y)^l \sum_{r=0}^{l+3} \binom{l+3}{r} K_1^r, \end{aligned}$$

by the binomial theorem. Since the binomial coefficients $\binom{l+3}{r}$ are sma- **192**
ller than 2^{l+3} for $r = 0, 1, \ldots, l+3$ and all $l = 0, 1, \ldots$, we can write for each term on the right

$$\sum_{r=0}^{l+3} \binom{l+3}{r} (\gamma_6 y)^l K_1^r \prec \sum_{r=0}^{l+3} 2^{l+3} (\gamma_6 y)^l K_1^r = 8 \sum_{r=0}^{l+3} (2\gamma_6 y)^l K_1^r.$$

Setting $l + r = h$, it follows that

$$\begin{aligned} \sum_{l=0}^{\infty} (\gamma_6 y)^l \sum_{r=0}^{l+3} \binom{l+3}{r} K_1^r &\prec 8 \sum_{h=0}^{\infty} \sum_{r=0}^{h} (2\gamma_6 y)^{h-r} K_1^r \\ &\prec 8 \sum_{h=0}^{\infty} \sum_{r=0}^{h} \binom{h}{r} (2\gamma_6 y)^{h-r} K_1^r = 8 \sum_{h=0}^{\infty} (2\gamma_6 y + K_1)^h, \end{aligned}$$

again by the binomial theorem, so that we have the majorization

$$K_1 \prec 8\gamma_7 y \sum_{h=0}^{\infty} (2\gamma_6 y + K_1)^h = \frac{8\gamma_7 y}{1 - (2\gamma_6 y + K_1)}. \tag{3.5.37}$$

Now let L be a power-series with indeterminate coefficients satisfying the algebraci equation

$$L = \frac{8\gamma_7 y}{1 - (2\gamma_6 y + L)}. \tag{3.5.38}$$

Once again, as in the case of K_1, one can use the implicit function theorem to obtain L. Since we are interested only in finding a majorant for K_1, we shall first show that L majorizes K_1 and then find a majorant for L itself. The coefficients of K_1 and L can be determined inductively by comparing coefficients of y^k, $k = 0, 1, \ldots,$ on both sides in (3.5.36) and
193 (3.5.38) respectively. It is clear that both K_1 and L lack constant terms and also that the coefficient of y in K_1 is γ_7, while in L it is $8\gamma_7$. Suppose that the coefficients of $y, \ldots, y^{k-1}$ in K_1 are majorized by those in L. Then by (3.5.37) we have

$$\{K_1\}_k \leq \left\{\frac{8\gamma_7 y}{1 - (2y_6 y + K_1)}\right\}_k = \left\{8y_7 y \sum_{h=0}^{\infty} (2\gamma_6 y + K_1)^h\right\}_k .$$

It is easy to see that the right side involves only the coefficients of $y, \ldots, y^{k-1}$ in K_1 and hence is smaller than

$$\left\{8\gamma_7 y \sum_{h=0}^{\infty} (2\gamma_6 y + L)^h\right\}_k = \left\{\frac{8\gamma_7 y}{1 - (2\gamma_6 y + L)}\right\}_k .$$

Then it follows by induction that $K_1 \prec L$. We majorize L further in the following way. Since

$$2\gamma_6 y \prec \frac{2\gamma_6 y}{1 - (2\gamma_6 y + L)}, \tag{3.5.39}$$

if we write $M = 2\gamma_6 y + L$, then by (3.5.38) and (3.5.39),

$$M = 2\gamma_6 y + L \prec \frac{(2\gamma_6 + 8\gamma_7)y}{1 - M} = \frac{\gamma_8 y}{1 - M}.$$

Let us denote by N the power series in y with undetermined coefficients satisfying the algebraic equation

$$N = \frac{\gamma_8 y}{1-N}. \tag{3.5.40}$$

Then it is easily seen as before that $M \prec N$. From (3.5.40) it follows that N satisfies the equation $N^2 - N + \gamma_8 y = 0$, which can also be written as **194**

$$4N^2 - 4N + 1 = 1 - 4\gamma_8 y, \text{ or, } (1-2N)^2 = 1 - 4\gamma_8 y.$$

From this we get

$$1 + 4N \prec 1 + 4N + \ldots = (1-2N)^2 = (1 - 4\gamma_8 y)^{-1},$$

and hence,

$$4N \prec (1 - 4\gamma_8 y)^{-1} - 1 = 4\gamma_8 y(1 - 4\gamma_8 y)^{-1}.$$

In other words, we have the majorization

$$K_1 \prec N \prec (1 - 4\gamma_8 y)^{-1}\gamma_8 y,$$

and the last is a geometric series in $4\gamma_8 y$ converging for $|y| < \dfrac{1}{4\gamma_8}$. Hence we conclude that the power-series $F_k(y_1, \ldots, y_n)$, $k = 1, \ldots, m$, converge for $|y_k| < \dfrac{1}{4\gamma_8}$, which completes the proof of the convergence. We have thus proved Theorem 3.5.1 under the restriction (3.5.19).

We shall now remove the restriction (3.5.19). The eigenvalues of A are again all real, distinct and non-zero but need no longer to satisfy the restriction (3.5.19). We shall show that in this case, the solutions $y_k(s)$ of (3.5.10), and hence the solutions $x_k(s)$ of (3.5.2), will be now power-series in the variables s, $e^{\lambda_1 s}, \ldots, e^{\lambda_n s}$, and not power series in the $e^{\lambda_k s}$, $k = 1, \ldots, n$, alone. We shall give only the construction of the solution and the proof of convergence will be on exactly the same lines **195**
as in the previous case.

First of all we remark that we can no longer use the relation (3.5.22):

$$c_k\left(\lambda_k - \sum_{l=1}^{n} g_l \lambda_l\right) = \left\{-\psi_l + \sum_{l=1}^{n} F_{ky_l}\psi_l\right\}_{g_1 \cdots g_n},$$

to determine the coefficients in the power-series F_k. If for a given $k = 1, \ldots, n$, we have

$$\lambda_k = \sum_{l=1}^{n} g_l \lambda_l, \tag{3.5.41}$$

for a finite set of n-tuples $(g_1, \ldots, g_n)$ of non-negative integers $g_1, \ldots, g_n$ with $g_1 + \ldots g_n = g \geq 2$, then unless the right side of (3.5.22) vanishes for each of these n-tuples, we would get a contradiction. Since the right side of (3.5.22) may not necessarily vanish for all such n-tuples $(g_1, \ldots, g_n)$ for which (3.5.41) holds, we cannot use the argument above for determining the coefficients c_k of $y_1^{g_1} \ldots y_n^{g_n}$ for this exceptional set $(g_1, \ldots, g_n)$. Hence it is not possible to determine the power-series F_k from the requirement (3.5.15), namely that $\chi(u_1, \ldots, u_n, 0, \ldots, 0) \equiv 0$ in $u_1, \ldots, u_n$. We therefore modify the proof in the following way.

We replace the requirement (3.5.15) by a weaker condition. We allow $\chi(u_1, \ldots, u_n, 0, \ldots, 0)$ to be a polynomial in $u_1, \ldots, u_n$ for just those k for which (3.5.41) holds. Let $V_k(u_1, \ldots, u_n)$ be polynomials in the n
196 variables $u_1, \ldots, u_n$, with real undetermined coefficients such that

$$\chi_k(u_1, \ldots, u_n, 0, \ldots, 0) = V_k(u_1, \ldots, u_n), \; k = 1, \ldots, m \tag{3.5.42}$$

Since $\lambda_k \neq \sum\limits_{l=1}^{n} g_l \lambda_l$ for $k = n+1, \ldots, m$, we may assume that

$$V_k(u_1, \ldots, u_n) \equiv 0, \; k = n+1, \ldots, m, \tag{3.5.43}$$

and consider (3.5.42) only for $k = 1, \ldots, n$. In this case we assume that every term in V_k is of the form $\alpha_k u_1^{g_1} \ldots u_n^{g_n}, g_1, \ldots, g_n$ non-negative integers with $g_1 + \ldots + g_n \geq 2$ for which (3.5.41) holds. There exist only finitely many such n-tuples $(g_1, \ldots, g_n)$, and hence only finitely many α_k, which determine the polynomial V_k. In order to determine α_k we observe that, since in (3.5.11) the power-series F_k start with quadratic terms, the coefficients of $y_1^{g_1} \ldots y_n^{g_n}$ in V_k, considered as a function of $y_1, \ldots, y_n$ after substituting (3.5.11) for $u_1, \ldots, u_n$, is precisely α_k and

$$\alpha_k u_1^{g_1} \ldots u_n^{g_n} = \alpha_k y_1^{g_1} \ldots y_n^{g_n} + \ldots.$$

Hence we can determine α_k by equating the coefficients of $y_1^{g_1} \ldots y_n^{g_n}$ on both sides of (3.5.42), considering V_k and χ_k as power-series in

$y_1, \ldots, y_n$. We define the coefficient c_k of $y_1^{g_1} \ldots y_n^{g_n}$ in F_k for which $g_1, \ldots, g_n$ satisfy (3.5.41) to be zero. Then we have

$$\begin{aligned} \alpha_k = \Big\{ &\psi_k(y_1, \ldots, y_n, F_{n+1}, \ldots, F_m) \\ &- \sum_{l=1}^{n} F_{ky_l}(y_1, \ldots, y_n)\psi_l(y_1, \ldots, y_n, F_{n+1}, \ldots, F_m) \\ &\quad + \alpha_k y_1^{g_1} \ldots y_n^{g_n} - V_k(y_1 - F_1, \ldots, y_n - F_n) \Big\}_{g_1 \ldots g_n} \end{aligned} \tag{3.5.44}$$

for all $g_1, \ldots, g_n$ satisfying (3.5.41) for the k in question. In (3.5.44) 197
α_k can be determined explicitly provided that the right side is known. The coefficients of $y_1^{g_1} \ldots y_n^{g_n}$ where $g_1, \ldots, g_n$ satisfy (3.5.41) are by assumption zero, while for all other $g_1, \ldots, g_n$ we have (3.5.19), so that the coefficients can be determined as before from (3.5.22) where now the term $\{V_k(y_1 - F_1, \ldots, y_n - F_n)\}_{g_1 \ldots g_n}$ has to be added on the right side. Then the polynomials V_k are completely determined. The convergence of V_k as power-series in $y_1, \ldots, y_n$ can be proved without much difficulty.

To obtain all solutions of (3.5.14): $u_k' = \lambda_k u_k + \chi_k$, $k = 1, \ldots, m$, we set as before, $v = \sum\limits_{k=n+1}^{m} u_k^2$, and then using (3.5.14), we have

$$v' = 2 \sum_{k=n+1}^{m} u_k u_k' = 2 \sum_{k=n+1}^{m} (\lambda_k u_k^2 + u_k \chi_k).$$

Since $V_k = 0$ for $k = n+1, \ldots, m$, and in each term of χ_k for such k we get one of $u_{n+1}, \ldots, u_m$ as a factor, our previous argument goes through. We have, as before, $v' \geq \lambda v$ and since $ve^{-\lambda s} \geq 0$ and nondecreasing, we have $v = 0$, so that $u_{n+1} = \ldots = u_m \equiv 0$. Substituting in (3.5.14) we obtain the system of differential equations

$$u_k' = \lambda_k u_k + V_k(u_1, \ldots, u_n), \; k = 1, \ldots, n,$$

where each V_k is a polynomial with real coefficients containing only terms of the form $u_1^{g_1} \ldots u_n^{g_n}$ where $g_1, \ldots, g_n$ satisfy (3.5.41). We arrange $\lambda_1, \ldots, \lambda_n$ in decreasing order and assume that

$$0 > \lambda_1 > \ldots > \lambda_n.$$

Then we claim that for every $k = 1, \ldots, n$, V_k is a polynomial only in the variables $u_1, \ldots, u_{k-1}$. To see this, consider the typical term $\alpha_k u_1^{g_1} \ldots u_n^{g_n}$ in V_k. We shall prove that $g_k = g_{k+1} = \ldots = g_n = 0$. Suppose, if possible, that for some $l, k+1 \leq l \leq n$, we have $g_l \neq 0$, so $g_l > 0$. Since (3.5.41) is satisfied, we have $-\lambda_k = \sum\limits_{r=1}^{n} g_r(-\lambda_r)$. But $0 < -\lambda_1 < \ldots < -\lambda_n$ by our ordering of the eigenvalues and so each $g_r(-\lambda_r) \geq 0$. Since $g_l > 0$ and $-\lambda_l > 0$, it follows that **198**

$$-\lambda_k \geq g_l(-\lambda_l) \geq -\lambda_l, \text{ or } \lambda_k \leq \lambda_l,$$

which is impossible since $l > k$. Hence $g_{k+1} = \ldots = g_n = 0$ necessarily, and it only remains to prove that $g_k = 0$. Suppose, if possible, that $g_k > 0$. Since $g_1 + \ldots + g_k = g_1 + \ldots + g_n \geq 2$, we have only two possibilities, either $g_k = 1$ or $g_k \geq 2$. If $g_k = 1$, then at least one of $g_1, \ldots, g_{k-1}$ is an integer ≥ 1 and hence $-\lambda_k > g_k(-\lambda_k) = -\lambda_k$, which is a contradiction. If $g_k \geq 2$, then $g_k(-\lambda_k) > -\lambda_k$, so $-\lambda_k > -\lambda_k$, which is again a contradiction. Hence $g_k = 0$.

So finally we obtain the following system of differential equations for $u_1, \ldots, u_m$:

$$\begin{aligned} u_k' &= \lambda_k u_k + V_k(u_1, \ldots, u_{k-1}),\ k = 1, \ldots, n, \\ u_k' &= \lambda_k u_k = 0,\ k = n+1, \ldots, m. \end{aligned} \tag{3.5.45}$$

199 We determine the general solution of (3.5.45) inductively. Since $V_1 \equiv 0$, $u_1' = \lambda_1 u_1$ and hence $u_1 = c_1 e^{\lambda_1 s}$, where c_1 is a constant of integration. Next, V_2 contains only terms of the form $\alpha_2 u_1^{g_1}$ where $g_1 \geq 2$ and $\lambda_2 = \lambda_1 g_1$. There is only one integral solution g_1 of $\lambda_2 = g_1 \lambda_1$, so that we have

$$V_2(u_1) = \alpha_2 u_1^{g_1}.$$

Inserting $u_1 = c_1 e^{\lambda_1 s}$ in $V_2(u_1)$, we get the differential equation

$$u_2' = \lambda_2 u_2 + \alpha_2 c_1^{g_1} e^{\lambda_1 g_1 s} = \lambda_2 u_2 + \alpha_2 c_1^{g_1} e^{\lambda_2 s},$$

which is the same as $(u_2 e^{-\lambda_2 s})' = V_2(c_1) = \alpha_2 c_1^{g_1}$, and on integration this gives

$$u_2 e^{-\lambda_2 s} = \alpha_2 c_1^{g_1} s + c_2, \text{ or } u_2 = (c_2 + V_2(c_1) s) e^{\lambda_2 s}.$$

If we denote the polynomial $V_2(c_1)s$ by $\mathscr{P}_2(c_1, s)$, then

$$u_2 = (c_2 + \mathscr{P}_2(c_1, s))e^{\lambda_2 s}.$$

Let us suppose that we have already proved that

$$u_{k-1} = (c_{k-1} + \mathscr{P}_{k-1}(c_1, c_2, \dots, c_{k-2}, s)e^{\lambda_{k-1} s},$$

where $\mathscr{P}_{k-1}$ is a polynomial in the $k-1$ real variables $c_1, \dots, c_{k-2}$ and s, and vanishes for $s = 0 : \mathscr{P}_{k-1}$ is uniquely determined by $0 = V_1, V_2, \dots, V_{k-1}$. Since $\mathscr{P}_1 \equiv 0$, we have seen that this holds for $k = 2, 3$. We now prove that

$$u_k = (c_k + \mathscr{P}_k(c_1, \dots, c_{k-1}, s)e^{\lambda_k s}.$$

We set **200**

$$c_l + \mathscr{P}_l(c_1, \dots, c_{l-1}, s) = Q_l(c_1, \dots, c_l, s), \; l = 1, \dots, k-1.$$

Then from (3.5.45), we have the differential equation

$$u'_k = \lambda_k u_k + V_k(Q_1 e^{\lambda_1 s}, \dots, Q_{k-1} e^{\lambda_{k-1}} s).$$

We recall once again that all the terms of V_k are of the form $\alpha_k u_1^{g_1} \dots u_{k-1}^{g_{k-1}}$, where $g_1, \dots, g_n$ satisfy the relations $\lambda_k = \sum\limits_{l=1}^{k-1} g_l \lambda_1$. Hence

$$V_k(Q_1 e^{\lambda_1 s}, \dots, Q_{k-1} e^{\lambda_{k-1} s}) = V_k(Q_1, \dots, Q_{k-1})e^{\lambda_k s},$$

so that we have the differential equation

$$u'_k = \lambda_k u_k + V_k(Q_1, \dots, Q_{k-1})e^{\lambda_k s},$$

$$\text{or,} \qquad (u_k e^{-\lambda_k s})' = V_k(Q_1, \dots, Q_{k-1}),$$

which, on integration from 0 to s gives,

$$u_k = (c_k + \mathscr{P}_k(c_1, \dots, c_{k-1}, s)e^{\lambda_k s}).$$

where $\mathscr{P}_k(c_1, \dots, c_{k-1}, s) = \int\limits_0^s V_k(Q_1, \dots, Q_{k-1})ds$. This proves our assertion. Here $c_1, \dots, c_k$ are constants of integration and are uniquely

determined by the initial values $u_k(0)$. We observe that since $V_k(Q_1, \ldots, Q_{k-1})$ is a polynomial in $c_1, \ldots, c_{k-1}, s$, and each contains a positive 201
power of s, it follows that $\mathscr{P}_k(c_1, \ldots, c_{k-1}, 0) = 0$. Hence $u_k(0) = c_k$, so that if the solutions $u_1, \ldots, u_n$ are to satisfy the relation $\sum_{k=1}^{m} u_k(s)^2 < \epsilon$, then we should necessarily have $\sum_{k=1}^{m} c_k^2 < \epsilon$. However, if $u_k(0) = c_k$ where $\sum_{k=1}^{m} c_k^2 < \epsilon$, then this may not imply that $\sum_{k=1}^{m} u_k(s)^2 < \epsilon$ for all $s \geq 0$. In the previous case, all the $\mathscr{P}_k$ were zero and $\lambda_1, \ldots, \lambda_n < 0$ and we had $\sum_{k=1}^{m} u_k(s)^2 < \epsilon$. But in the present case, this is not in general true for all $s \geq 0$. However, since $u_k(s) = Q_k(c_1, \ldots, c_{k-1}, s)e^{\lambda_k s}$ and Q_k is a polynomial while $\lambda_k < 0$, it follows that for sufficiently large s, $u_k(s)$ are so small that $\sum_{k=1}^{m} u_k(s)^2 < \epsilon$, and moreover, $u_k(s) \to 0$ as $s \to \infty$. This again is a constructive method of determining the solutions.

In order to obtain the solution of the original system of equations in the unknown functions x_k, $k = 1, \ldots, m$, we first solve for $y_1, \ldots, y_m$ in terms of $u_1, \ldots, u_m$. We have, by inversion of (3.5.11),

$$
\begin{aligned}
y_k &= u_k + G_k(u_1, \ldots, u_n),\ k = 1, \ldots, n,\\
y_k &= G_k(u_1, \ldots, u_m),\ k = n+1, \ldots, m,
\end{aligned}
$$

where G_k are power-series with real coefficients, starting with quadratic terms. Hence the $y_l (l = 1, \ldots, m)$ are power-series in $(c_k + \mathscr{P}_k(c_1, \ldots, c_{k-1}, s)e^{\lambda_k s}, k = 1, \ldots, n$. Since the x_l are linear functions of y_l, the same
202 assertion holds for x_l also and thus we obtain all the asymptotic solutions of the original system $x' = Ax + \varphi(x)$. They involve n real parameters $c_1, \ldots, c_n$. This completes the proof of Theorem 3.5.1. □

We shall now consider the situation in which the eigenvalues of the matrix $A = (a_{kl})$ are not necessarily real. We have

Theorem 3.5.2. *Suppose that the eigen-values of the matrix $A = (a_{kl})$ are distinct, some possibly complex, and that all eigen-values have non-*

zero real parts. Then the general solutions $x_k(s)$ in $s \geq 0$ of the system

$$x'_k = \sum_{l=1}^{m} a_{kl} x_l + \varphi_k(x_1, \ldots, x_m), \ k = 1, \ldots, m.$$

which satisfy the condition $\sum\limits_{k=1}^{m} x_k(s)^2 < \epsilon$ for small $\epsilon > 0$, involve as many real parameters as the number of eigen-values with negative real parts.

Proof. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigen-values of the real matrix $A = (a_{kl})$, and let $\lambda_k = \rho_k + i\tau_k$, $k = 1, \ldots, m$. Since A is real, its characteristic polynomial has real coefficients and so its complex roots occur in pairs of complex conjugates. Hence if λ_k is a complex eigen-value, the conjugate complex $\overline{\lambda}_k$ is also an eigen-value λ_l, where $l = l_k$; so $\lambda_l = \overline{\lambda}_k$ for $l = l_k$. Then $\overline{\lambda}_l = \lambda_k$ and so $l_{l_k} = k$ for all k. If λ_k is a real eigen-value, then $\lambda_k = \overline{\lambda}_k = \lambda_l$ and since the eigen-values are simple, $l = l_k = k$. Hence $(l_1, \ldots, l_m)$ is a permutation of $(1, \ldots, m)$ and since the λ_k are simple, the permutation consists entirely of transpositions. □

We consider a linear transformations $x = Cy$, C being a complex **203**
matrix with $|C| \neq 0$. Since all the eigen-values of A are distinct and different from zero, we can find a C such that $C^{-1}AC = D$ is in the normal diagonal form:

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}.$$

Then $AC = CD$ and we can determine the matrix C from this condition. Let C_k, $k = 1, \ldots, m$, denote the columns of C. Then we have $AC_k = \lambda_k C_k$, $k = 1, \ldots, m$, and this can be seen immediately by comparing the elements on both sides of $AC = CD$. Hence C_k is an eigen-vector of the matrix A belonging to the eigen-value λ_k. Since the λ_k are distinct, the eigen-vectors C_k are all distinct. These eigen-vectors are uniquely determined up to constant scalar factors, in general complex. Again using our earlier notation,

$$A\overline{C}_k = \overline{A}\,\overline{C}_k = \overline{\lambda}_k \overline{C}_k = \lambda_l \overline{C}_k, \ l = l_k,$$

so that $\overline{C_k}$ is an eigen-vector belonging to the eigen-value λ_l where $l = l_k$. Hence $\overline{C_k}$ is a scalar multiple of C_l and so by a suitable normalization we may assume that

$$\overline{C_k} = C_l,\ \overline{C_l} = C_k,\ l = l_k,\ k = 1, \ldots, m. \tag{3.5.46}$$

Since the matrix C is complex, y is a complex vector. Since x is real, $x = \overline{x}$ and so $Cy = \overline{C}\overline{y}$. It follows from this, by (3.5.46), that

$$C_k y_k + C_l y_l = \overline{C_k}\overline{y_k} + \overline{C_l}\overline{y_l} = C_l\overline{y_k} + C_k\overline{y_l}.$$

204 But $y = C^{-1}x$ is a uniquely determined vector and so we should have

$$y_l = \overline{y}_k,\ y_k = \overline{y}_l,\ l = l_k,\ k = 1, \ldots, m.$$

As we have to deal with formal power-series, we shall drop the assumption that x is real. The relation $x = Cy = \overline{C}\overline{y}$ can be given a sense even when x is not real if we define formally the indeterminates $\overline{y}_1, \ldots, \overline{y}_m$ by setting

$$y_l = \overline{y}_k,\ y_k = \overline{y}_l,\ l = l_k,\ k = 1, \ldots, m. \tag{3.5.47}$$

(Thus $\overline{y}_1, \ldots, \overline{y}_m$ is just a permutation of the indeterminates $y_1, \ldots, y_m$, this permutation consisting entirely of transpositions).

By the substitution $x = Cy$, the given system of differential equations $x' = Ax + \varphi(x)$ goes over into the system

$$y' = Dy + C^{-1}\varphi(x), \tag{3.5.48}$$

where $C^{-1}\varphi(x)$ is a column vector $\sigma(x)$ of power-series $\sigma_k(x)$. The coefficients of σ_k are complex and they are obtained in the following way. If α_k denotes the coefficient of $x_1^{g_1} \ldots x_m^{g_m}$ of degree $g = g_1 + \ldots + g_m \geq 2$ in $\varphi_k(x)$, then the coefficient of $x_1^{g_1} \ldots x_m^{g_m}$ in $\sigma_k(x)$ is given by $\sum\limits_{l=1}^{m} d_{kl}\alpha_l$, $C^{-1} = (d_{kl})$.

205 If $f = f(x_1, \ldots, x_m)$ is a formal power-series with complex coefficients in the m indeterminates $x_1, \ldots, x_m$, then we denote by $\overline{f} = \overline{f}(x_1, \ldots, x_m)$ the power-series obtained by replacing the coefficients in

f by their complex conjugates and by f^* that obtained by replacing the coefficients in f by their absolute values: thus

$$\{\bar{f}\}_{g_1\ldots g_m} = \overline{\{f\}_{g_1\ldots g_m}} \text{ and } \{f^*\}_{g_1\ldots g_m} = |\{f\}_{g_1\ldots g_m}|.$$

Because of $\bar{\varphi}(x) = \varphi(x) = C\sigma(x)$ we have $\bar{\sigma}_k(x) = \sigma_l(x)$. Since $Cy = x = \bar{C}\bar{y}$, we have also $\bar{\sigma}_k(\bar{C}\bar{y}) = \sigma_l(Cy)$. Denoting $\sigma_k(Cy)$ by $\psi_k(y)$, we have, by the last formula, $\psi_l(y) = \bar{\psi}_k(\bar{y})$. Hence we can rewrite the system of differential equations (3.5.48) in the form

$$y' = Dy + \psi(y), \tag{3.5.49}$$

where $\psi(y)$ is a column-vector whose components are power-series $\psi_k(y)$ with complex coefficients and starting with quadratic terms.

Let $\lambda_1, \ldots, \lambda_n$ denote the eigen-values whose real parts are negative and $\lambda_{n+1}, \ldots, \lambda_m$ those whose real parts are positive; so

$$\rho_1 < 0, \ldots, \rho_n < 0;\ \rho_{n+1} > 0, \ldots, \rho_m > 0. \tag{3.5.50}$$

Since for $l = l_k$, $\lambda_l = \bar{\lambda}_k = \rho_k - i\tau_k$, it follows that as the index k runs through $1, \ldots, n$, l_k also runs through $1, \ldots, n$, and if k runs through $n+1, \ldots, m$, so does l_k.

As in the proof of Theorem 3.4.1, we now make a non-linear transformation of the variables $y_1, \ldots, y_m$ to the variables $u_1, \ldots, u_m$ of the form 206

$$u_k = y_k - F_k(y_1, \ldots, y_n),\ k = 1, \ldots, m,$$

where the F_k are power-series in the variables $y_1, \ldots, y_n$ alone, with complex coefficients and starting with quadratic terms. As in the case of real eigen-values, the differential equations (3.5.49) are transformed into

$$u_k' = \lambda_k u_k + \chi_k(u_1, \ldots, u_m),\ k = 1, \ldots, m, \tag{3.5.51}$$

where $\chi_k(u_1, \ldots, u_m)$ is defined by

$$\chi_k = \lambda_k F_k + \psi_k - \sum_{l=1}^{n} F_{ky_l}(\lambda_l y_l + \psi_l). \tag{$*$}$$

Our argument for the construction of the power-series F_k goes through as in the real case.

First suppose that $\lambda_k \neq \sum\limits_{l=1}^{n} g_l \lambda_l$ for all n-tuples of non-negative integers $g_1, \ldots, g_n$ with $g = g_1 + \ldots + g_n \geq 2$. This is only a finite set of conditions on the λ_k, as in the real case. We determine F_k again by requiring that $\chi_k(u_1, \ldots, u_n, 0, \ldots, 0) \equiv 0$, i.e.

$$\begin{aligned} \lambda_k F_k + \psi_k(y_1, \ldots, y_n,\ F_{n+1}, \ldots, F_m) \\ - \sum_{l=1}^{n} F_{ky_l}(\lambda_l y_l + \psi_l(y_1, \ldots, \\ y_n,\ F_{n+1}, \ldots, F_m)) = 0. \end{aligned} \tag{$**$}$$

$k = 1, \ldots, m$, and we get as before a recurrence formula for finding the coefficients in F_k. If $c_k y_1^{g_1} \ldots y_n^{g_n}$ is a term of total degree $g = g_1 + \ldots g_n \geq 2$ in F_k, then comparing the coefficients of $y_1^{g_1} \ldots y_n^{g_n}$, we have

$$c_k \left(\lambda_k - \sum_{l=1}^{n} g_l \lambda_1 \right) = \{ -\psi_k(y_1, \ldots, y_n,\ F_{n+1}, \ldots, F_m) + \sum_{l=1}^{n} F_{ky_1} \psi_l \}_{g_1 \ldots g_n}.$$

207 We find c_k by induction as in the real case. The convergence of the power-series F_k thus obtained is proved as in the real case. Let $\alpha = \min(-\rho_1, \ldots, -\rho_n) > 0$. Then there exists a positive number γ_1 depending only on α, such that

$$|\lambda_k - \sum_{l=1}^{n} g_l \lambda_l| \geq |\operatorname{Re}(\lambda_k - \sum_{l=1}^{n} g_l \lambda_l)| > \gamma_1^{-1}(g_1 + \ldots + g_n).$$

On the other hand, the ψ_k are power-series convergent in a complex neighbourhood $|y_1| < \underline{\rho_1}, \ldots, |y_m| < \underline{\rho_m}$ of $y_1 = 0, \ldots, y_m = 0$. We shall denote by $\gamma_2, \gamma_3 \ldots$ large positive constants. If $|\psi_k| \leq \gamma_2$ in this neighbourhood, then by Cauchy's theorem,

$$|\{\psi_k\}_{h_1 \ldots h_m}| \leq \gamma_2 \underline{\rho_1}^{-h_1} \ldots \underline{\rho_m}^{-h_m},$$

for all n-tuples of non-negative integers $h_1, \ldots, h_m$ with $h_1 + \ldots + h_m \geq 2$. So we have

$$\psi_k, \prec \sum_{h_1 + \ldots + h_m \geq 2} \gamma_2 \frac{y_1^{h_1} \ldots y_m^{h_m}}{\underline{\rho}_1^{h_1} \ldots \underline{\rho}_m^{h_m}} \prec \sum_{h=2}^{\infty} \gamma_2 \left(\frac{y_1}{\underline{\rho}_1} + \ldots + \frac{y_m}{\underline{\rho}_m} \right)^h$$

$$\prec \gamma_2\gamma_3^2 \frac{(y_1 + \ldots + y_m)^2}{1 - \gamma_3(y_1 + \ldots + y_m)} \equiv \Psi(y_1, \ldots, y_m),$$

where $\gamma_3^{-1} = \min(\underline{\rho}_1, \ldots, \underline{\rho}_m) > 0$. Then we have

$$|c_k|(g_1 + \ldots + g_n) \leq \gamma_1|\{\psi_k + \sum_{l=1}^{n} F_{ky_1}\psi_l\}_{g_1 \ldots g_n}|.$$

We replace F_k by F_k^* and ψ_k, ψ_l by Ψ and we see that $F_k \prec F_k^*$ and **208**

$$|c_k|(g_1 + \ldots + g_n) \leq |\gamma_4 \left\{(1 + \sum_{l=1}^{n} F_{ky_l}^*)\Psi\right\}_{g_1 \ldots g_n}|.$$

The rest of the proof is the same as before and we conclude that the power-series $F_k, k = 1, \ldots, m$, converge for complex values of $y_1, \ldots, y_n$ in a complex neighbourhood of $y_1 = 0, \ldots, y_n = 0$. (The F_k can now be regarded as convergent power-series in the complex variables $y_1, \ldots, y_m$).

We can now obtain y_k in terms of u_k by locally inverting the substitution $u_k = y_k - F_k(y_1, \ldots, y_n)$, $k = 1, \ldots, m$, and we find that

$$y_k = u_k + G_k(u_1, \ldots, u_m), \ k = 1, \ldots, m, \tag{3.5.52}$$

where the G_k are uniquely determined power-series with complex coefficients. Since the F_k converge as functions of the complex variables $y_1, \ldots, y_n$, we can now look upon $u_1, \ldots, u_m$ also as complex variables and G_k are therefore convergent power-series in the complex variables $u_1, \ldots, u_m$. Since we are interested in real solutions $x_k(s)$ of the original system, we have to find out under what conditions for $u_1, \ldots, u_m$ we shall have $\overline{y}_k = y_l$, $l = l_k$, $k = 1, \ldots, m$, so that $x_1, \ldots, x_m$ all are real. We might conjecture that this condition is again $\overline{u}_k = u_l$, $l = l_k$, $k = 1, \ldots, m$, and conversely. This is true and it is enough to prove this in the following formal situation. Suppose we introduce the indeterminates $\overline{y}_k = y_l$, $l = l_k$; we shall prove that if $u_1, \ldots, u_m$ and $\overline{u}_1, \ldots, \overline{u}_m$ are defined by $u_k = y_k - F_k(y_1, \ldots, y_n)$, $\overline{u}_k = \overline{y}_k - \overline{F}_k(\overline{y}_1, \ldots, \overline{y}_n)$ then $\overline{u}_k = u_l$, $l = l_k$, $k = 1, \ldots, m$. And for this purpose it is enough to prove **209** that $\bar{F}_k(\bar{y}_1, \ldots, \bar{y})_n = F_l(y_1, \ldots, y_n)$, or equivalently, that

$$\bar{F}_k(y_{l_1}, \ldots, y_{l_n}) = F_{l_k}(y_1, \ldots, y_n), \; k = 1, \ldots, m.$$

For this we observe that the condition $\chi_k(u_1, \ldots, u_n, \; 0, \ldots, 0) \equiv 0$, $k = 1, \ldots, m$, gives the following identity in the formal power-series:

$$\begin{aligned}\lambda_{l_k} F_{l_k}(y_1, \ldots, y_n) - \sum_{r=1}^{n} \lambda_{l_r} y_{l_r} F_{l_r} y_{l_r}(y_1, \ldots, y_n) \qquad (3.5.53)\\ = -\psi_{l_k}(y_1, \ldots, y_n, F_{n+1}, \ldots, F_m) + \sum_{r=1}^{n} F_{l_k} y_{l_r}(y_1, \ldots, y_n)\\ \psi_{l_r}(y_1, \ldots, y_n, \; F_{n+1}, \ldots, F_m).\end{aligned}$$

By replacing the coefficients on both sides by their complex conjugates we have the identity

$$\begin{aligned}\bar{\lambda}_{l_k} \bar{F}_{l_k}(y_1, \ldots, y_n) - \sum_{r=1}^{n} \bar{\lambda}_{l_r} y_{l_r} \bar{F}_{l_k y_{l_r}}(y_1, \ldots, y_n) \qquad (3.5.53)'\\ = -\bar{\psi}_{l_k}(y_1, \ldots, y_n, \bar{F}_{n+1}, \ldots, \bar{F}_m)\\ + \sum_{r=1}^{n} \bar{F}_{l_k y_{l_r}}(y_1, \ldots, y_n) \bar{\psi}_{l_r}(y_1, \ldots, y_n, \bar{F}_{n+1}, \ldots, \bar{F}_m).\end{aligned}$$

(Recall that $\bar{f}$ denotes the power-series whose coefficients are the complex conjugates of those of f). The permutation $y_{l_1}, \ldots, y_{l_m}$ of the indeterminates $y_1, \ldots, y_m$ introduces the indeterminates $\bar{y}_1, \ldots, \bar{y}_m$. We have seen that since $l_{l_k} = k$, $k = 1, \ldots, m$, $\psi_k(\underline{y_1}, \ldots, y_m) = \bar{\psi}_l(\bar{y}_1, \ldots, \bar{y}_m) = \bar{\psi}_l(y_{l_1}, \ldots, y_{l_m})$. Now using the fact that $\bar{\lambda}_{l_k} = \lambda_k$ and $\bar{y}_{l_r} = y_r$, we get from (3.5.53)′ the following identity:

$$\begin{aligned}\lambda_k \bar{F}_{l_k}(\bar{y}_1, \ldots, \bar{y}_n) - \sum_{r=1}^{n} \lambda_r y_r \bar{F}_{l_k y_r}(\bar{y}_1, \ldots, \bar{y}_n)\\ = -\psi_l(y_1, \ldots, y_n, \bar{F}_{l_{n+1}}(\bar{y}_1, \ldots, \bar{y}_n), \ldots, \bar{F}_{l_m}(\bar{y}_1, \ldots, \bar{y}_n))\\ + \sum_{r=1}^{n} \bar{F}_{l_k y_r}(\bar{y}_1, \ldots, \bar{y}_n) \psi_r(y_1, \ldots, y_n, \bar{F}_{l_{n+1}}\end{aligned}$$

$$(\bar{y}_1, \ldots, \bar{y}_n), \ldots, \bar{F}_{l_m}(\bar{y}_1, \ldots, \bar{y}_n)) \tag{3.5.54}$$

because the passage from $y_1, \ldots, y_m$ to $\bar{y}_1, \ldots, \bar{y}_m$ is nothing but the permutation $y_{l_1}, \ldots, y_{l_m}$, and moreover, when k runs through $1, \ldots, n$, so does l_k and when k runs through $n+1, \ldots, m$, so does l_k. Hence $\bar{F}_l(\bar{y}_1, \ldots, \bar{y}_n)$, $l = l_k$, $k = 1, \ldots, m$ satisfy the functional equation $(**)$ for $F_k(y_1, \ldots, y_n)$. On the other hand, $F_k(y_1, \ldots, y_n)$ is uniquely determined by $(**)$. Hence by the uniquenes of the solution of $(*)$, we have, for $l = l_k$, $k = 1, \ldots, m$, $\bar{F}_l(\bar{y}_1, \ldots, \bar{y}_n) = F_k(y_1, \ldots, y_n)$. It is also clear that if F_k converges as a power series in the complex variables $y_1, \ldots, y_n$ in a complex neighbourhood of $y_1 = 0, \ldots, y_n = 0$ and we replace $y_1, \ldots, y_n$ in $\bar{F}_k(y_1, \ldots, y_n)$ by the complex conjugates $\bar{y}_1, \ldots, \bar{y}_n$, then the corresponding variable u_k defined by $u_k = y_k - F_k(y_1, \ldots, y_n)$ goes over into the conjugate variable $\bar{u}_k$. This proves the assertion that $\bar{u}_k = \bar{u}_l$, $l = l_k$, $k = 1, \ldots, m$. 210

Now it follows from the definition $(*)$ that also

$$\bar{\chi}_k(\bar{u}_1, \ldots, \bar{u}_m) = \chi_l(u_1, \ldots, u_m).$$

Thus we have the reduced system of differential equations

$$u'_k = \lambda_k u_k + \chi_k, \ u_l = \bar{u}_k, \ l = l_k, \ k = 1, \ldots, m,$$

under the assumption that $\sum\limits_{k=1}^{m} |u_k|^2 < \epsilon$ for all $s \geq 0$ and sufficiently small ϵ. In order to obtain the explicit solutions we consider the function $v(s)$ defined by 211

$$v = \sum_{k=n+1}^{m} |u_k|^2 = \sum_{k=n+1}^{m} u_k \bar{u}_k.$$

Differentiating with respect to s we have

$$v' = \sum_{k=n+1}^{m} (u'_k \bar{u}_k + u_k \bar{u}'_k).$$

It is clear that v' is real; since $\lambda_l = \bar{\lambda}_k$, $u_l = \bar{u}_k$, $\chi_l(u) = \bar{\chi}_k(\bar{u})$ for $l = l_k$, we see that $\bar{u}'_k = u'_l = \lambda_l u_l + \chi_l(u) = \bar{\lambda}_k \bar{u}_k + \bar{\chi}_k(\bar{u})$. Hence

$$v' = \sum_{k=n+1}^{m} (\lambda_k + \bar{\lambda}_k) u_k \bar{u}_k + \sum_{k=n+1}^{m} (u_k \bar{\chi}_k(\bar{u}) + \bar{u}_k \chi_k(u))$$

$$= 2\sum_{k=n+1}^{m} \rho_k |u_k|^2 + 2\sum_{k=n+1}^{m} \mathrm{Re}(\overline{u}_k \chi_k(u)).$$

Let $\rho = \min(\rho_{n+1}, \ldots, \rho_m) > 0$. Since χ_k starts with quadratic terms and each term of χ_k contains at least one of $u_{n+1}, \ldots, u_m$ as a factor, and χ_k is uniformly convergent, we show as in the real case that for sufficiently small ϵ,

$$2\sum_{k=n+1}^{m} \mathrm{Re}(\overline{u}_k \chi_k(u)) \geq -\rho v.$$

212 So we have $v' \geq \rho v$ and hence $(ve^{-\rho s})' \geq 0$ and $ve^{-\rho s}$ is non-decreasing. Since $ve^{-\rho s} \geq 0$ it follows that $v = 0$, or $u_k(s) = 0$ for all $s \geq 0$, $k = n+1, \ldots, m$. Thus the system of differential equations is further reduced to

$$u'_k = \lambda_k u_k,\ k = 1, \ldots, n.$$

Hence

$$u'_l = \lambda_l u_l,\ l = l_k,\ k = 1, \ldots, n, \overline{\lambda}_k = \lambda_l.$$

Integrating these we obtain

$$u_k = c_k e^{\lambda_k s},\ u_l = c_l e^{\lambda_l s}, l = l_k, k = 1, \ldots, n.$$

Since $\lambda_l = \overline{\lambda}_k$ and $u_l = \overline{u}_k$ for $l = l_k$, $k = 1, \ldots, n$, and all $s \geq 0$, we have also $c_l = \overline{c}_k$. Thus we get exactly n real parameters in the real solution $x_k = x_k(s)$ of the original system asymototic to the equilibrium solution. This proves Theorem 3.4.2 under the restriction $\lambda_k \neq \sum\limits_{l=1}^{n} \lambda_l g_l$.

It is easy to extend the argument to the case in which $\lambda_k = \sum\limits_{l=1}^{n} \lambda_l g_l$ for some given k, by imposing the same condition as in the real case, namely, $\chi_k(u_1, \ldots, u_n,\ 0, \ldots, 0) = V_k(u_1, \ldots, u_n)$, a polynomial with complex coefficients in $u_1, \ldots, u_n$ consisting entirely of terms of the form $\alpha_k u_1^{g_1} \ldots u_n^{g_n}$ where $g_1, \ldots, g_n$ are non-negative integers with $g_1 + \ldots + g_n \geq 2$, satisfying $\lambda_k = \sum\limits_{l=1}^{n} \lambda_l g_l$. Once again, if $0 < \rho_1 < \ldots < \rho_n$, we can show that V_k is actually a polynomial in $u_1, \ldots, u_{k-1}$ only, $\sum\limits_{k=1}^{m} |u_k|^2 <$

213 ϵ for ϵ sufficiently small. We have now, in addition, $\overline{V}_l(\overline{u}_1, \ldots, \overline{u}_n) = V_k(u_1, \ldots, u_n)$, $l = l_k$. Then the solution are given by

$$u_k = Q_k(c_1, \ldots, c_k, s)e^{\lambda_k s}, k = 1, \ldots, n;$$
$$u_k = 0, \; k = n+1, \ldots, m,$$

where Q_k are polynomials in s with complex coefficients, $q_k(c_1, \ldots, c_k, 0) = c_k$ and $c_l = \overline{c}_k$, $\overline{Q}_k(\overline{c}_1, \ldots, \overline{c}_k, s) = Q_l(c_1, \ldots, c_k, s)$. Once again, the general solution asymptotic to the equilibrium solution contains n real parameters, and this completes the proof of Theorem 3.5.2.

It remains now to consider the solutions in the case in which the matrix $A = (a_{kl})$ may have multiple eigen-values. The problem is more complicated in this case, the difficulty being that the matrix A cannot now be reduced to the diagonal form. We can nevertheless extend our earlier results to this case. We have

Theorem 3.5.3. *Suppose that the matrix A has eigen-values which are, in general, complex, all with non-zero real parts, some of them possibly multiple. Then the general solution of the system $x' = Ax+\varphi(x)$ such that $\sum\limits_{k=1}^{m} x_k^2 < \epsilon$ for small $\epsilon > 0$ contains exactly n arbitrary real parameters, n being the number of eigen-values with negative real parts.*

Proof. We transform the matrix A in the following way. We can find a matrix C with $|C| \neq 0$ such that $C^{-1}AC = D$ breaks up into boxes along the main diagonal. More precisely, the matrix $D = (d_{kl})$ has the following property: if $\lambda_1, \ldots, \lambda_m$ are the eigen-values of A, 214

$$d_{kk} = \lambda_k, d_{kl} = 0 \text{ for } k \neq l \text{ and } k \neq l+1, \tag{3.5.55}$$

$d_{kl} = 0$ for $k = l+1$ if $\lambda_l \neq \lambda_{l+1}$, $d_{kl} = 0$ or 1 for $k = l+1$ if $\lambda_l = \lambda_{l+1}$.

Let $e_1 = 0$, $e_k = 0$ if $\lambda_k \neq \lambda_{k-1}$, $e_k = 0$ or 1 if $\lambda_k = \lambda_{k-1}$, $k = 2, \ldots, m$. Then the transformed system $y' = Dy + \psi(y_1, \ldots, y_m)$ can be written in the form

$$y_k' = \lambda_k y_k + e_k y_{k-1} + \psi_k(y_1, \ldots, y_m), \; k = 1, \ldots, m, \tag{3.5.56}$$

where ψ_k again are power-series with complex coefficients starting with quadratic terms and converging in a neighbourhood of $y_1 = 0, \ldots, y_m = 0$. Let $\lambda_k = \rho_k + i\tau_k$ and let us suppose that

$$0 < -\rho_1 < -\rho_2 < \ldots < -\rho_n; \rho_k > 0,\ k = n+1, \ldots, m, \qquad (3.5.57)$$

so that $\lambda_1, \ldots, \lambda_n$ are all the eigen-values with negative real parts. As before we perform the non-linear transformation

$$u_k = y_k - F_k(y_1, \ldots, y_n),\ k = 1, \ldots, m, \qquad (3.5.58)$$

where F_k are power-series in $y_1, \ldots, y_n$ with complex coefficients to be determined, starting with quadratic terms. We restrict ourselves to the case in which $\lambda_k \neq \sum\limits_{r=1}^{n} g_r\lambda_r$, $k = 1, \ldots, m$, for non-negative integers $g_1, \ldots, g_n$ with $g_1 + \ldots g_n \geq 2$. When this is not the case the proof can be modified as in the earlier situations.

215 From (3.5.58), by differentiation, on using (3.5.56), we have

$$u'_k = \lambda_k(u_k + F_k) + e_k y_{k-1} + \psi_k - \sum_{r=1}^{n} F_{ky_r}(\lambda_r y_r + e_r y_{r-1} + \psi_r),$$

which can be written in the form

$$u'_k = \lambda_k u_k + e_k u_{k-1} + \chi_k(u_1, \ldots, u_m), k = 1, \ldots, m,$$

where

$$\chi_k = \lambda_k F_k - \sum_{r=1}^{n} \lambda_r y_r F_{ky_r} + e_k F_{k-1} - \sum_{r=1}^{n} e_r y_{r-1} F_{ky_r} + \psi_k - \sum_{r=1}^{n} \psi_r F_{ky_r}. \qquad (3.5.59)$$

Once again we determine the power-series F_k by requiring that

$$\chi_k(u_1, \ldots, u_n, 0, \ldots, 0) = 0,\ k = 1, \ldots, m.$$

Since $u_k = 0$ implies that $y_k = F_k(y_1, \ldots, y_n)$, $k = n+1, \ldots, m$, this condition implies the following identity:

$$\lambda_k F_k - \sum_{r=1}^{n} \lambda_r y_r F_{ky_r} = -e_k F_{k-1} + \sum_{r=1}^{n} e_r y_{r-1} F_{ky_r}$$

$$- \psi_k(y_1, \dots, y_n, F_{n+1}, \dots, F_m) + \sum_{r=1}^{n} \psi_r(y_1, \dots, y_n, F_{n+1}, \dots, F_m) F_{ky_r}. \tag{3.5.60}$$

Let $c_{k,g_1,\dots,g_n} y_1^{g_1} \dots y_n^{g_n}$ denote a term of total degree $g = g_1 + \dots + g_n \geq 2$ in the power-series F_k. Comparing the coefficients of $y_1^{g_1} \dots y_n^{g_n}$ on both sides of (3.5.60), we have

$$k, g_1, \dots, g_n \left(\lambda_k - \sum_{r=1}^{n} g_r \lambda_r \right) = \{-\psi_k + \sum_{r=1}^{n} F_{ky_r} \psi_r\}_{g_1,\dots,g_n} - e_k c_{k-1,g_1,\dots,g_n} + \\ + \sum_{r=1}^{n} e_r (g_r + 1) c_{k,g_1,\dots,g_{r-1}-1,\ g_r+1,\dots,g_n}. \tag{3.5.61}$$

Since coefficients of terms of total degree g, namely $c_{k-1,g_1,\dots,g_n}$ and 216
$c_{k,g_1,\dots,g_{r-1}-1,g_r+1,\dots,g_n}$, $r = 1, \dots, n$, occur on the right side of (3.5.60), we cannot carry out the induction construction for the coefficients as in the previous cases. However, we can argue by induction on introducing a lexicographic ordering for the indices in the subscripts.

First of all, suppose that the coefficients $c_{k-1,g_1,\dots,g_n}$ of total degree $g_1 + \dots + g_n = g$ have already been determined. Then the second term on the righ side of (3.5.61) is known. For $k = 1$ the term corresponding to this does not appear and so in order to determine $c_{1,g_1,\dots,g_n}$ it is enough to consider only the terms

$$\left\{ -\psi_1 + \sum_{r=1}^{n} \psi_r F_{1y_r} \right\}_{g_1,\dots,g_n} + \sum_{r=1}^{n} e_r (g_r + 1) c_{1,g_1,\dots g_{r-1}-1,g_r+1,\dots,g_n}.$$

If all the coefficients of terms of total degrees $2, \dots, g - 1$ have been determined in $F_1, \dots, F_n$, then since $\{-\psi_1 + \sum\limits_{r=1}^{n} \psi_r F_{1y_r}\}_{g_1,\dots,g_n}$ involves only these coefficients, it follows that it is known. Thus we have to deal only with the term

$$\sum_{r=1}^{n} e_r (g_r + 1) c_{1,g_1,\dots,g_{r-1}-1,g_r+1,\dots,g_n}.$$

For this we introduce the lexicographic ordering for the subscripts $(g_1, \ldots, g_n)$. If $(g_1, \ldots, g_n)$ and $(h_1, \ldots, h_n)$ are n-tuples of non-negative integers, then we say that $(g_1, \ldots, g_n)$ is lower than $(h_1, \ldots, h_n)$, written $(g_1, \ldots, g_n) < (h_1, \ldots, h_n)$ if the first of the non-vanishing differences $g_1 - h_1, \ldots, g_n - h_n$ is negative. It is clear that this ordering is transitive: if $(g_1, \ldots, g_n) < (h_1, \ldots, h_n)$ and $(h_1, \ldots, h_n) < (k_1, \ldots, k_n)$, then $(g_1, \ldots, g_n) < (k_1, \ldots, k_n)$. In this ordering we find that each set $(g_1, \ldots, g_{r-1} - 1, g_r + 1, \ldots, g_n)$ of the subscripts of $c_{k,g_1,\ldots,g_{r-1}-1,g_r+1,\ldots,g_n}$, $r = 2, \ldots, n$, is lower than $(g_1, \ldots, g_n)$. (Since $e_1 = 0$ by definition, the case $r = 1$ is taken care of). We carry out the induction in the following manner. Given $(g_1, \ldots, g_n)$, suppose that we have already determined 217

(i) all coefficients in $F_1, \ldots, F_n$ of total degrees $2, \ldots, g - 1$;

(ii) the coefficient $c_{k-1,g_1,\ldots,g_n}$ of $y_1^{g_1} \ldots y_n^{g_n}$ in F_{k-1}; and

(iii) all the coefficients $c_{k,h_1,\ldots,h_n}$ of $y_1^{h_1} \ldots y_n^{h_n}$ in F_k of total degree $h_1 + \ldots + h_n = g$ where $(h_1, \ldots, h_n) < (g_1, \ldots, g_n)$.

Then we can determine the coefficient $c_{k,g_1,\ldots,g_n}$ of $y_1^{g_1} \ldots y_n^{g_n}$ in F_k from the recurrence formula (3.5.61).

If $k = 1$ and $(g_1, \ldots, g_n) = (0, \ldots, 0, 2)$, then the coefficient $c_{1,0,\ldots,0,2}$ is determined by the coefficient of y_n^2 in ψ_1 and so is known. Then we can determine all the coefficients of the quadratic terms in $F_k, k = 1, \ldots, n$,
218 successively from the recurrence relation (3.5.61). Hence we can begin the induction. Thus, whenever $\lambda_k \neq \sum_{r=1}^{n} g_r \lambda_r$, $k = 1, \ldots, m$, all the coefficients in $F_1, \ldots, F_n$ can be determined by induction on the lexicographic ordering of the subscripts. If this condition is not satisfied for some k, then we set, as before, $\chi_k(u, \ldots, u_n, 0, \ldots, 0) = V_k(u_1, \ldots, u_n)$, where V_k are polynomials consisting only of terms of the form $\alpha_k u_1^{g_1} \ldots u_n^{g_n}$ with $\lambda_k = \sum_{r=1}^{r} g_r \lambda_r$. The proof is easily modified to suit this case and we shall not go into the details.

Before discussing the condition in order that the solution $x_k(s)$ be real, we shall prove the convergence of the power-series F_k obtained above. This presents some difficulty; the estimates we obtained in the case of simple eigen-values do not suffice. The coefficients $c_{k,g_1,\ldots,g_n}$ are determined from the recurrence formula (3.5.61) and we have assumed that $\lambda_k - \sum\limits_{r=1}^{n} g_r\lambda_r \neq 0$, $k = 1,\ldots,m$. We have

$$|\lambda_k - \sum_{r=1}^{n} g_r\lambda_r| \geq \rho_k - \sum_{r=1}^{n} g_r\rho_r| \geq \rho_k - \sum_{r=1}^{n} g_r\rho_r,\ k = 1,\ldots,m.$$

If $\alpha = \min(-\rho_1,\ldots,-\rho_n) > 0$, we can write

$$|\lambda_k - \sum_{r=1}^{n} g_r\lambda_r| \geq \rho_k + \alpha(g_1+\ldots+g_n) = \rho_k + \frac{\alpha}{2}(g_1+\ldots+g_n-1) + \frac{\alpha}{2}(g_1+\ldots+g_n+1).$$

There are at most finitely many n-tuples $(g_1,\ldots,g_n)$ for which $\rho_k + \dfrac{\alpha}{2}(g_1+\ldots+g_n-1) < 0$, $k = 1,\ldots,m$, $g_1+\ldots+g_n \geq 2$. For all other n-tuples we have $|\lambda_k - \sum\limits_{r=1}^{n} g_r\lambda_r| \geq \dfrac{\alpha}{2}(g_1+\ldots+g_n+1)$. Since $|\lambda_k - \sum\limits_{r=1}^{n} g_r\lambda_r| > 0$, we can find a constant $\gamma_1 \geq 1$ such that $|\lambda_k - \sum\limits_{r=1}^{n} g_r\lambda_r| >$ **219**
$\gamma_1^{-1}(g_1+\ldots+g_n+1)$, $k = 1,\ldots,m$ for all n-tuples $(g_1,\ldots,g_n)$. Hence, from (3.5.61),

$$\begin{aligned}
&\gamma_1^{-1}(g_1+\ldots+g_n+1)|c_{k,g_1,\ldots,g_n}| \leq e_k|c_{k-1,g_1,\ldots,g_n}| + \\
&+ \sum_{r=2}^{n} e_r(g_r+1)|c_{k,g_1,\ldots,g_{r-1}-1,g_r+1,\ldots,g_n}| + |\{-\psi_k + \sum_{r-1}^{n} F_{ky_r}\psi_r\}_{g_1,\ldots,g_n}|.
\end{aligned} \tag{3.5.62}$$

Denoting by F_k^* the formal power-series obtained by replacing the coefficients in F_k by their absolute values, we have $F_k \prec F_k^*$, $k = 1,\ldots,m$. We denote by $\gamma_2,\gamma_3,\ldots$ sufficiently large positive constants. As in the case of real eigen-values we have the majorization for ψ_k:

$$\psi_k \prec \sum_{g=2}^{\infty} \gamma_2^g(y_1+\ldots y_m)^g = \frac{\gamma_2^2(y_1+\ldots+y_m)^2}{1-\gamma_2(y_1+\ldots+y_m)}.$$

We define

$$\Psi(y_1,\ldots,y_n) = \frac{\gamma_2^2(y_1+\ldots+y_n+F_{n+1}^*(y_1,\ldots,y_n)+\ldots+F_m^*(y_1,\ldots,y_n))^2}{1-\gamma_2(y_1+\ldots+y_n+F_{n+1}^*(y_1,\ldots,y_n)+\ldots+F_m^*(y_1,\ldots,y_n))}$$

and we have

$$\psi_k(y_1,\ldots,y_n,F_{n+1}(y_1,\ldots,y_n),\ldots,F_m(y_1,\ldots,y_n)) \prec \Psi(y_1,\ldots,y_n).$$

We next observe that the coefficient of $y_1^{g_1}\ldots y_n^{g_n}$ in F_k^* is $|c_{k,g_1,\ldots,g_n}|$,
220 while its coefficient in $y_r F^*_{ky_r}$ is $g_r|c_{k,g_1,\ldots,g_n}|$, so that

$$(1+g_1+\ldots+g_n)|c_{k,g_1,\ldots,g_n}| = \left\{F_k^* + \sum_{r=1}^{n} y_r F^*_{ky_r}\right\}_{g_1,\ldots,g_n}$$

Now (3.4.62) implies the majorization

$$\gamma_1^{-1}\left(F_k^* + \sum_{r=1}^{n} y_r F^*_{ky_r}\right) \prec e_k F^*_{k-1} + \sum_{r=2}^{n} e_r y_{r-1} F^*_{ky_r} + \left(1+\sum_{r=1}^{n} F^*_{ky_r}\right)\Psi. \tag{3.5.63}$$

From this we want to obtain a majorization similar to what we had in the case of simple eigen-values. In order to achieve this we replace all the F_k^* by just one function F, which majorizes all the F_k^*, independently of k. Let $\mu_1,\ldots,\mu_n$ be positive numbers ≤ 1 which we shall choose suitably later. Let $F = \sum\limits_{k=1}^{m} \mu_k F_k^*$, so that $F_k \prec F_k^* \prec \mu_k^{-1} F$, $k = 1,\ldots,m$. It would thus be sufficient to prove the convergence of the power-series F, with non-negative coefficients, starting with quadratic terms. Multiplying both sides of (3.5.63) by μ_k and summing over $k = 1,\ldots,m$, and making use of the facts that $\sum\limits_{k=1}^{m} \mu_k F^*_{ky_r} = F_{y_r}$, $e_1 = 0$ and $e_k \leq 1$, $k = 2,\ldots,n$, we have

$$\begin{aligned}\gamma_1^{-1}\Big(F + \sum_{r=1}^{n} y_r F_{y_r}\Big) &\prec \sum_{k=2}^{m} \mu_k F^*_{k-1} + \sum_{r=2}^{m} e_r y_{r-1} F_{y_r} + \sum_{k=1}^{n} \mu_k \Psi + \sum_{r=1}^{n} F_{y_r}\Psi \\ &\prec \sum_{k=2}^{m} \mu_k F^*_{k-1} + \sum_{r=2}^{n} y_{r-1} F_{y_r} + \left(m + \sum_{r=1}^{n} F_{y_r}\right)\Psi. \end{aligned} \tag{3.5.64}$$

We now choose μ_k in such a way that all the ratios μ_k/μ_{k-1} are indepen-
221 dent of k: let $\mu_1 = 1$, $\mu_k/\mu_{k-1} = \gamma_1^{-1}$, $k = 2,\ldots,m$, then all $\mu_k \leq 1$. Since

$\mu_k = \dfrac{\mu_k}{\mu_{k-1}} \cdot \mu_{k-1} = \gamma_1^{-1}\mu_{k-1}$, $k = 2, \ldots, m$, we see that

$$\sum_{k=2}^{m} \mu_k F^*_{k-1} = \gamma_1^{-1} \sum_{k=2}^{m} \mu_{k-1} F^*_{k-1} \prec \gamma_1^{-1} F.$$

Substituting this in (3.5.64) we have

$$\gamma_1^{-1} \sum_{r=1}^{n} y_r F_{y_r} \prec \sum_{r=2}^{n} y_{r-1} F_{y_r} + (m + \sum_{r=1}^{n} F_{v_r})\Psi.$$

Once again we use the fact that each $F^*_k \prec \mu_k^{-1} F$ to obtain a majorant for Ψ. From the definition of Ψ we have

$$\begin{aligned}\Psi(y_1, \ldots, y_m) &= \sum_{g=2}^{\infty} \gamma_2^g (y_1 + \ldots + y_n + F^*_{n+1} + \ldots + F^*_m)^g \\ &\prec \sum_{g=2}^{\infty} \gamma_2^g (y_1 + \ldots + y_n + (\mu_{n+1}^{-1} + \ldots + \mu_m^{-1})F)^g.\end{aligned}$$

Taking $\gamma_3 = \mu_{n+1}^{-1} + \ldots + \mu_m^{-1}$, we have

$$y_1 + \ldots + y_n + F^*_{n+1} + \ldots + F^*_m \prec \gamma_3(y_1 + \ldots + y_n + F),$$

and hence, with $\gamma_4 = \gamma_2\gamma_3$,

$$m\Psi(y_1, \ldots, y_m) \prec m \sum_{g=2}^{\infty} \gamma_4^g (y_1 + \ldots y_n + F)^g \equiv \Phi, \text{ say.}$$

Hence we have

$$\gamma_1^{-1} \sum_{r=1}^{n} y_r F_{y_r} \prec \sum_{r=2}^{n} y_{r-1} F_{y_r} + (1 + \sum_{r=1}^{n} F_{y_r})\Phi. \qquad (3.5.65)$$

Let $\nu_1, \ldots, \nu_n$ be positive numbers, to be chosen suitably later. Setting $y_k = \nu_k y$, $k = 1, \ldots, n$, and $F(y_1, \ldots, y_n) = F(\nu_1 y, \ldots, \nu_n y) = G(y)$, we have 222

$$G_y = \sum_{r=1}^{n} F_{y_r} \frac{dy_r}{dy} = \sum_{r=1}^{n} \nu_r F_{y_r},$$

so that $yG_y = \sum\limits_{r=1}^{n} \nu_r y F_{y_r} = \sum\limits_{r=1}^{n} y_r F_{y_r}$. Then from (3.5.65) we have

$$\gamma_1^{-1} y G_y \prec \sum_{r=2}^{n} y_{r-1} F_{y_r} + \left(1 + \sum_{r=1}^{n} F_{y_r}\right) \Phi.$$

Since, by definition, $y_{r-1}/y_r = \nu_{r-1}/\nu_r$, $y_{r-1}F_{y_r} = y_r F_{y_r} \nu_{r-1}/\nu_r$, $r = 2, \ldots, n$ we obtain

$$\sum_{r=2}^{n} y_{r-1} F_{y_r} = \sum_{r=2}^{n} \frac{\nu_{r-1}}{\nu_r} y_r F_{y_r}.$$

Now choose $\nu_1 = 1$ and ν_{r-1}/ν_r to be independent of r, say, $\nu_{r-1}/\nu_r = (2\gamma_1)^{-1}$, $r = 2, \ldots, n$. Then

$$\sum_{r=2}^{n} y_{r-1} F_{y_r} \prec (2\gamma_1)^{-1} \sum_{r=1}^{n} y_r F_{y_r} = (2\gamma_1)^{-1} y G_y.$$

On the other hand, since $\nu_r = 2\gamma_1 \nu_{r-1} = (2\gamma_1)^{r-1} \nu_1 \geq 1$, we have $F_{y_r} \prec \nu_r F_{y_r}$, so that we have

$$\sum_{r=1}^{n} F_{y_r} \prec \sum_{r=1}^{n} \nu_r F_{y_r} = G_y,$$

which implies that $1 + \sum\limits_{r=1}^{n} F_{y_r} \prec 1 + G_y$. Let $\gamma_5 = m\gamma_4(\nu_1 + \ldots + \nu_n)$ and

$$\wedge \equiv \sum_{g=2}^{\infty} \gamma_5^g (y + G)^g = \frac{\gamma_5^2 (y + G)^2}{1 - \gamma_5 (y + G)}.$$

Then it is clear that $\Phi \prec \wedge$. Thus finally we obtain the majorization

$$\gamma_1^{-1} y G_y \prec (2\gamma_1)^1 y G_y + (1 + G_y) \wedge,$$

223 or,

$$yG_y \prec 2\gamma_1(1+g_y)\wedge .$$

This is of the same form as the majorization we had for the power-series H in the case of simple eigen-values. We proceed exactly as in that case and prove that $G(y)$ converges for $|y| < \delta$ for sufficiently small positive δ. Then it follows that F_k^* and F_k converge for $y_1, \ldots, y_n$ in the complex region $|y_1| < \nu_1\delta, \ldots, |y_n| < \nu_n\delta$. We have thus proved the convergence of F_k in a complex neighbourhood of $(0, \ldots, 0)$.

We shall inidicate briefly the construction of F_k and the proof of its convergence in the case in which for some k, $\lambda_k = \sum\limits_{r=1}^{n} g_r\lambda_r$. As we have mentioned already, the F_k in this case are so chosen that $\chi_k(u_1, \ldots, u_n, 0, \ldots, 0) = V_k(u_1, \ldots, u_n)$, $k = 1, \ldots, m$, where V_k is a polynomial with complex coefficients containing only terms of the form $\alpha_{k,g_1,\ldots,g_n} u_1^{g_1} \ldots u_n^{g_n}$ for those $g_1, \ldots, g_n$ with $g_1 + \ldots + g_n \geq 2$ which satisfy $\lambda_k = \sum\limits_{r=1}^{n} \lambda_r g_r$. (We recall that there are only finitely many such n-tuples). In F_k we define the coefficient $c_{k,g_1,\ldots,g_n}$ to be zero whenever $\lambda_k = \sum\limits_{r=1}^{n} g_r\lambda_r$. Substituting $u_k = y_k - F_k(y_1, \ldots, y_n)$ in the polynomials V_k, we can express V_k as a power-series in the complex variables $y_1, \ldots, y_n$ with complex coefficients and starting with quadratic terms. Also, since F_k start with quadratic terms, we have

$$u_l^{g_l} = (y_l - F_l(y_1, \ldots, y_n))^{g_1} = y_l^{g_1} + \text{ terms of degree } > g_l,$$

and hence $\alpha_{k,g_1,\ldots,g_n} u_1^{g_1} \ldots u_n^{g_n} = \alpha_{k,g_1,\ldots,g_n} y_1^{g_1} \ldots y_n^{g_n} +$ terms of degree $>$ **224**
$g_1 + \ldots + g_n$. Now substituting $u_l = y_l - F_l(y_1, \ldots, y_n)$, $l = 1, \ldots, n$ in the condition $\chi_k(u_1, \ldots, u_n, 0, \ldots, 0) = V_k(u_1, \ldots, u_n)$ and comparing coefficients of $y_1^{g_1} \ldots y_n^{g_n}$ on both sides, we find that

$$\alpha_{k,g_1,\ldots,g_n} = \left\{\psi_k - \sum_{r=1}^{n} F_{ky_r}\psi_r\right\}_{g_1,\ldots,g_n} + e_k c_{k-1,g_1,\ldots,g_n}$$
$$- \sum_{r=2}^{n} e_r(g_r+1)c_{k,g_1,\ldots,g_{r-1}-1,\, g_r+1,\ldots,g_n} + \quad \text{a polynomial in the}$$

coefficients in $F_1, \ldots, F_n$ and $V_1, \ldots, V_n$ of degrees $\leq g-1$. We determine from this recurrence formula all the coefficients $\alpha_{k,g_1,\ldots,g_n}$ in V_k by induction on the lexicographic ordering of the subscripts $g_1, \ldots, g_n$ and the natural ordering of k; the coefficients $c_{k,h_1,\ldots,h_n}$ are determined from (3.5.61) in the same way when $\lambda_k \neq \sum\limits_{l=1}^{n} g_l \lambda_l$. There remains the proof of the convergence of F_k and V_k as power-series in $y_1, \ldots, y_n$ which is trivial for the polynomial V_k. Since

$$V_k(u_1, \ldots, u_n) = \sum_{g_1,\ldots,g_n,\ \lambda_k = \sum g_l \lambda_l} \alpha_{k,g_1,\ldots,g_n} u_1^{g_1} \ldots u_n^{g_n},$$

if $\gamma_6 = \max |\alpha_{k,g_1,\ldots,g_n}|$, then we have

$$V_k(u_1, \ldots, u_n) \prec \gamma_6 \sum_{h_1+\ldots+h_n \geq 2} u_1^{h_1} \ldots u_n^{h_m} \prec \gamma_6 \sum_{h=2}^{\infty} (u_1 + \ldots + u_n)^h.$$

225 Since $u_l = y_l - F_l(y_1, \ldots, y_n)$ and so $u_l \prec y_l + F_l^*$, we have

$$V_k(u_1, \ldots, u_n) \prec \gamma_6 \sum_{h=2}^{\infty} (y_1 + \ldots + y_n + F_1^* + \ldots + F_n^*)^h.$$

If we define $F = \sum\limits_{k=1}^{n} \mu_k F_k^*$, μ_k positive constants as before, we have $F_k^* \prec \mu_k^{-1} F$, $k = 1, \ldots, n$, and

$$V_k \prec \gamma_6 \sum_{h=2}^{\infty} \gamma_7^h (y_1 + \ldots + y_n + F)^h \prec \frac{\gamma_8 (y_1 + \ldots + y_n + F)^2}{1 - \gamma_6 (y_1 + \ldots + y_n + F)},$$

which is of the same form as Φ and again we get a majorization of the type

$$yG_y \prec 2\gamma_1 (1 + G_y) \wedge .$$

The proof of the convergence of F_k proceeds in the same way.

We consider next the problem of finding a condition in order that the solution $x_k = x_k(s)$ of $x' = Ax + \varphi(x)$ be real. When all the eigen-values were simple we found that $\bar{\lambda}_k = \lambda_l$ for a uniquely determined $l = l_k$ and in that case it was enough to prove that

$$\bar{F}_k(\bar{y}_1, \ldots, \bar{y}_n) = F_l(y_1, \ldots, y_n), l = l_k, k = 1, \ldots, m,$$

where $\overline{y}_k = u_l$. In the case in which λ_k has multiplicity > 1, $\overline{\lambda}_k$ also has the (same) multiplicity > 1, A being a real matrix. It is clear that $l = l_k$ is not uniquely determined for those k for which λ_k has multiplicity > 1. However, we can rearrange the eigen-values λ_k and define l_k in such a way that for each k, l_k is uniquely determined. We shall illustrate this in 226
the case in which $e_k = 1$. Then, by (3.5.56), $\lambda_k = \lambda_{k-1}$. We now arrange the λ_k in such a way that $\overline{\lambda}_k = \lambda_l$ and $\overline{\lambda}_{k-1} = \lambda_{l-1}$; so that $\lambda_l = \lambda_{l-1}$. This means that we now define l_k, $k = 1, \ldots, n$, in such a way that if for some k, $\lambda_k = \lambda_{k-1}$, then $l_{k-1} = l_k - 1$.

Since x is real, $x = Cy = \overline{C}\overline{y}$. If C_k denotes the columns of the matrix C, we can choose again $C_l = \overline{C}_k$, $C_k = \overline{C}_l$ where $l = l_k$ corresponds to the above ordering of λ_k, and hence

$$C_k y_k + C_l y_l = C_l \overline{y}_k + C_k \overline{y}_l, \ k, l = 1, \ldots, m,$$

so that once again $\overline{y}_k = y_l$, where $l = l_k$ is uniquely defined as above. Then we apply the method used in the case of simple eigen-values to prove that

$$\overline{F}_k(\overline{y}_1, \ldots, \overline{y}_n) = \overline{F}_k(y_{l_1}, \ldots, y_{l_n}) = F_l(y_1, \ldots, y_n), l = l_k.$$

Formally, if we write $\overline{u}_k = u_l$, $l = l_k$, after the proof of the convergence of F_k in a complex neighbourhood of $y_1 = 0, \ldots, y_n = 0$, it would follow that u_l is the complex conjugate of the complex variable u_k when $l = l_k$ defined uniquely as above.

Now we proceed to find explicitly the solutions of

$$u'_k = \lambda_k u_k + e_k u_{k-1} + \chi_k(u_1, \ldots, u_m), \ k = 1, \ldots, m.$$

For this we apply a method similar to the one used in the case of simple eigen-values. Given a sufficiently small $\epsilon > 0$, if we seek a complex solution u_k such that $\sum\limits_{k=1}^{m} |u_k|^2 < \epsilon$, we prove that $u_{n+1} = \ldots = u_m = 0$. 227
Let $h_{n+1}, \ldots, h_m$ be $m-n$ positive constants, to be chosen later. Consider the positive function $v = v(s)$ defined for $s \geq 0$ by

$$v = \sum_{k=n+1}^{m} h_k |u_k|^2 = \sum_{k=n+1}^{m} h_k u_k \overline{u}_k.$$

Suppose that $0 \le v(s) < \epsilon$ for all $s \ge 0$. Then

$$|u_k| \le \sqrt{\frac{v(s)}{h_k}} < \sqrt{\frac{\epsilon}{h_k}},\ k = n+1, \dots, m.$$

We have $v' = \sum_{k=n+1}^{m} h_k(u_k' \bar{u}_k + u_k \bar{u}_k')$. From the differential equations we have

$$u_k' = \lambda_k u_k + e_k u_{k-1} + \chi_k(u),\ \bar{u}_k' = \bar{\lambda}_k \bar{u}_k + e_k \bar{u}_{k-1} + \bar{\chi}_k(\bar{u}).$$

So

$$\begin{aligned} v' &= \sum_{k=n+1}^{m} h_k(\bar{u}_k(\lambda_k u_k + e_k u_{k-1} + \chi_k(u)) + u_k(\bar{\lambda}_k \bar{u}_k + e_k \bar{u}_{k-1} + \bar{\chi}_k(\bar{u}))) \\ &= 2 \sum_{k=n+1}^{m} h_k \rho_k u_k \bar{u}_k + \sum_{k=n+1}^{m} h_k e_k (u_k \bar{u}_{k-1} + \bar{u}_k u_{l-1}) \\ &\quad + \sum_{k=n+1}^{m} h_k (u_k \bar{\chi}_k(\bar{u}) + \bar{u}_k \chi_k(u)). \end{aligned}$$

On the other hand,

$$\begin{aligned} &\sum_{k=n+1}^{m} h_k \rho_k u_k \bar{u}_k + \sum_{k=n+1}^{m} h_k e_k (u_k \bar{u}_{k-1} + \bar{u}_k u_{k-1}) \\ &\quad = \sum_{k=n+1}^{m} h_k \rho_k (u_k + \frac{e_k}{\rho_k} u_{k-1})(\bar{u}_k + \frac{e_k}{\rho_k} \bar{u}_{k-1}) - \sum_{k=n+1}^{m} \frac{h_k}{\rho_k} e_k^2 u_{k-1} \bar{u}_{k-1}. \end{aligned}$$

Since $e_k^2 = e_k$ and $e_{n+1} = 0$ (because λ_n has negative real part and λ_{n+1} positive real part, so that $\lambda_n \ne \lambda_{n+1}$), we have

$$\begin{aligned} \sum_{k=n+1}^{m} \frac{h_k}{\rho_k} e_k^2 u_{k-1} \bar{u}_{k-1} &= \sum_{k=n+1}^{m} \frac{h_k}{\rho_k} e_k u_{k-1} \bar{u}_{k-1} = \sum_{k=n}^{m-1} \frac{h_{k+1}}{\rho_{k+1}} e_{k+1} u_k \bar{u}_k \\ &= \sum_{k=n+1}^{m-1} \frac{h_{k+1}}{\rho_{k+1}} e_{k+1} u_k \bar{u}_k. \end{aligned}$$

228 Hence we obtain

$$
\begin{aligned}
v' = & \sum_{k=n+1}^{m} h_k\rho_k u_k\bar{u}_k + \sum_{h=n+1}^{m} h_k\rho_k\Big(u_k + \frac{e_k}{\rho_k}u_{k-1}\Big)\Big(\bar{u}_k + \frac{e_k}{\rho_k}\bar{u}_{k-1}\Big)- \\
& - \sum_{k=n+1}^{m} \frac{h_{k+1}}{\rho_{k+1}} e_{k+1}u_k\bar{u}_k + \sum_{k=n+1}^{m} h_k(u_k\bar{\chi}_k(\bar{u}) + \bar{u}_k\chi_k(u)) \\
& \geq \sum_{k=n+1}^{m-1} \Big(h_k\rho_k - \frac{h_{k+1}}{\rho_{k+1}}e_{k+1}\Big)u_k\bar{u}_k + h_m\rho_m u_m\bar{u}_m \\
& \qquad + \sum_{k=n+1}^{m} h_k(u_k\bar{\chi}_k(\bar{u}) + \bar{u}_k\chi_k(u)).
\end{aligned}
$$

Now we choose $h_{k+1} = \frac{1}{2}h_k\rho_k\rho_{k+1}$, $k = n+1, \ldots, m-1$, $h_{n+1} = 1$, so that $h_k\rho_k - \frac{h_{k+1}}{\rho_{k+1}}e_{k+1} = \frac{1}{2}h_k\rho_k$ when $e_{k+1} = 1$ and $= h_k\rho_k$ when $e_{k+1} = 0$. In any case we have $h_k\rho_k - \frac{h_{k+1}}{\rho_{k+1}}e_{k+1} \geq \frac{1}{2}h_k\rho_k$ ans also $h_m\rho_m > \frac{1}{2}h_m\rho_m$ (as $h_m\rho_m$ is positive). Hence

$$v' \geq \frac{1}{2}\sum_{k=n+1}^{m} h_k\rho_k u_k\bar{u}_k + \sum_{k=n+1}^{m} h_k(u_k\bar{\chi}_k(\bar{u}) + \bar{u}_k\chi_k(u)).$$

Let $\beta = \min(\rho_{n+1}, \ldots, \rho_m) > 0$. Then we have

$$v' \geq \frac{1}{2}\beta v + \sum_{k=n+1}^{m} h_k(u_k\bar{\chi}_k(\bar{u}) + \bar{u}_k\chi_k(u)).$$

As in our previous discussion, since $\chi_k(u_1, \ldots, u_n, 0, \ldots, 0) = 0$ implies that each term of $\chi_k(u_1, \ldots, u_m)$ contains at least one of $u_{n+1}, \ldots, u_m$ as a factor and χ_k starts with quadratic terms, and since $|u_k| < \sqrt{\frac{\epsilon}{h_1}}$, and the χ_k are uniformly convergent, we have, for sufficiently large s.

$$\Big|\sum_{k=n+1}^{m} h_k(u_k\bar{\chi}_k(\bar{u}) + \bar{u}_k\chi_k(u))\Big| \leq \frac{1}{4}\beta v.$$

Thus we obtain the differential inequality 229

$$v' \geq \frac{1}{4}\beta v, \ (ve^{-1/4\beta s})' \geq 0,$$

which implies that $ve^{-\frac{1}{4}\beta s}$ is non-decreasing, and since it is non - negative, we necessarily have $v = 0$. This means that

$$u_{n+1} = \ldots = u_m = 0.$$

The system of differential equations is therefore reduced to

$$u'_k = \lambda_k u_k + e_k u_{k-1}, k = 1, \ldots, n; u'_k = 0, \ k = n+1, \ldots, m,$$

in case $\lambda_k \neq \sum\limits_{r=1}^{n} g_r \lambda_r$ and to

$$\begin{aligned} &u'_k = \lambda_k u_k + e_k u_{k-1} + V_k(u_1, \ldots, u_{k-1}), \\ &k = 1, \ldots, n; u'_k = 0, \ k = n+1, \ldots, m, \end{aligned}$$

in the contrary case. Consider the first case. Since $e_1 = 0$, we have $u'_1 = \lambda_1 u_1$, or $(u_1 e^{-\lambda_1 s})' = 0$, which on integration from 0 to s gives

$$u_1 = c_1 e^{\lambda_1 s}, \ c_1 = u_1(0).$$

We insert this in $u'_2 = \lambda_2 u_2 + e_2 u_1$ where $e_2 = 0$, if $\lambda_2 \neq \lambda_1$ and $e_2 = 0$ or 1, if $\lambda_2 = \lambda_1$ and integrate from 0 to s. We can continue this procedure
230 and obtain all the u_k. In the second case, when for some k, $\lambda_k = \sum\limits_{r=1}^{n} g_r \lambda_r$, once again since $e_1 = 0$ and $V_1 = 0$, we have $u'_1 = \lambda_1 u_1$ from we obtain $u_1 = e_1 e^{\lambda_1 s}$. Suppose that we have already proved that

$$u_r = (c_r + \mathscr{P}_r(c_1, \ldots, c_{r-1}, s)e^{\lambda_r s} \equiv Q_r(c_1, \ldots, c_r, s)e^{\lambda_r s}),$$

for $r = 1, \ldots, k-1$, where $c_r = u_r(0)$ and $\mathscr{P}_r$ is a polynomial in $c_1, \ldots, c_{r-1}$ and s, which vanishes for $s = 0$. Then we show that

$$u_k = (c_k + \mathscr{P}_k(c_1, \ldots, c_{k-1}, s))e^{\lambda_k s} \equiv Q_k(c_1, \ldots, c_k, s)e^{\lambda_k s}.$$

If $e_k = 0$, $u'_k = \lambda_k u_k + V_k(u_1, \ldots, u_{k-1})$ and inserting $u_1, \ldots, u_{k-1}$ already found, and integrating from 0 to s, we get u_k as above, as in the case of simple eigen-values. If $e_k = 1$, then necessarily $\lambda_{k-1} = \lambda_k$ and hence

$$\begin{aligned} u'_k &= \lambda_k u_k + e_k u_{k-1} + V_k(u_1, \ldots, u_{k-1}) \\ &= \lambda_k u_k + Q_{k-1}(c_1, \ldots, c_{k-1}, s) e^{\lambda_{k-1} s} + V_k(u_1, \ldots, u_{k-1}) \\ &= \lambda_k u_k + Q_{k-1}(c_1, \ldots, c_{k-1}, s) e^{\lambda_k s} + V_k(Q_1, \ldots, Q_{k-1}) e^{\lambda_k s}, \end{aligned}$$

and we can integrate this to obtain u_k. Thus all the u_k are determined by induction.

As in the case of simple eigenvalues, it follows that if $c_k = u_k(0)$, then $\sum\limits_{k=1}^{m} |c_k|^2 < \epsilon$. However, if $\sum\limits_{k=1}^{m} |c_k|^2 < \epsilon$, it may no longer be true that $\sum\limits_{k=1}^{m} |u_k(s)|^2 < \epsilon$ for all $s \geq 0$, because of the presence of the term 231 $\mathscr{P}_k(c_1, \ldots, c_{k-1}, s)$ in u_k. However, $u_k(s) \to 0$ as $s \to \infty$.

Finally, if we arrange the eigen-values ans define l_k, $k = 1, \ldots, m$, in such a way that $\lambda_{k-1} = \lambda_k$, $l_{k-1} = l_k - 1$, and so $\overline{\lambda}_{k-1} = \lambda_{l-1}$, $\overline{\lambda}_k = \lambda_l$, then we can prove as in the case of simple eigen-values that $\overline{u}_k = u_l$, $l = l_k$, $k = 1, \ldots, m$. Then it follows that $\overline{c}_k = c_l$, $l = l_k$, $k = 1, \ldots, n$, and therefore a general solution such that $\sum\limits_{k=1}^{m} |u_k|^2 < \epsilon$ contains exactly n real parameters. This completes the proof of Theorem 3.5.3.

6 Application to the three-body problem

We shall now apply the general theory of stability of solutions of systems of ordinary differential equations to the special case of the system of differential equations of the three-body problem near a general collision.

We obtained in § 4 the system of differential equations (3.5.36) near the general collision at $t = 0$ or, equivalently, at $s = \infty$, where $s = e^{-t}$:

$$\delta'_k = \sum_{l=1}^{6} a_{kl} \delta_l + \varphi_k(\delta_1, \ldots, \delta_6), \ k = 1, \ldots, 6. \qquad (3.6.1)$$

where a_{kl} are real constants determined uniquely by the masses and the φ_k are power-series in $\delta_1, \ldots, \delta_6$ with real coefficients, also determined uniquely by the masses, starting with quadratic terms and convergent for small values of $|\delta_1|, \ldots, |\delta_6|$. We consider the corresponding normalized
232 system of differential equations in the unknown functions $u_k(s)$, $k = 1, \ldots, 6$, and shall the general theory of § 5 of solutions asymptotic to the equilibrium solution.

We have to discuss the two cases, the equilateral case and the collinear case, and we had computed, at the end of § 4, the eigen-values $\lambda_1, \ldots, \lambda_6$ of the matrix $A = (a_{kl})$ in both the cases. In the equilateral case there are three negative eigen-values:

$$\lambda_1 = -a_2 = \frac{1}{6} - \frac{1}{6}\sqrt{13 - 12\sqrt{1 - 3a}},$$
$$\lambda_2 = -a_1 = \frac{1}{6} - \frac{1}{6}\sqrt{13 + 12\sqrt{1 - 3a}},\ \lambda_3 = -a_o = -\frac{2}{3}$$

where $a = (m_1 m_2 + m_2 m_3 + m_1 m_3)\ (m_1 + m_2 + m_3)^{-2}$; $0 < a \leq \frac{1}{3}$ and $0 < -\lambda_1 \leq -\lambda_2 < -\lambda_3$, $\lambda_1 = \lambda_2$ only when $a = \frac{1}{3}$, or $m_1 = m_2 = m_3$. In the collinear case, there are two negative eigen-values:

$$\lambda_1 = -b_o = -\frac{2}{3}, \lambda_2 = -b_1 = \frac{1}{6} - \frac{1}{6}\sqrt{25 + 16b},$$
$$\text{where } b = \frac{m_1(1 + (1-\omega)^{-1} + (1-\omega)^{-2}) + m_3(1 + \omega^{-1} + \omega^{-2})}{m_1 + m_2(\omega^{-2} + (1-\omega)^2) + m_3}$$
$$> 0; 0 < -\lambda, < -\lambda_2.$$

We first consider the case in which $\lambda_k \neq \sum_{l=1}^{p} g_l \lambda_l$ for non-negative integers $g_1, \ldots, g_p$ with $g_1 + \ldots + g_p \geq 2$ ($p = 2$ in the collinear case and $p = 3$ in the equilateral case). In the equilateral case,

$$u_1 = c_1 e^{-a_2 s}, u_2 = c_2 e^{-a_1 s}, u_3 = c_3 e^{-a_o s}, \tag{3.6.2}$$

provided that not all m_1, m_2, m_3 are equal, and in the collinear case,

$$u_1 = c_1 e^{-b_o s},\ u_2 = c_2 e^{-b_1 s}. \tag{3.6.3}$$

233 From (3.6.2) and (3.6.3) we find that the essential difference between the equilateral and collinear cases consists in the fact that the general solution asymptotic to the equilibrium solution as $s \to \infty$ contain 3 real parameters in the former case and only 2 in the latter. We observe that a_1, a_2 and b_1 are in general irrational, whereas $a_o = b_o = \frac{2}{3}$ is rational. Since $e^{-s} = t$, we have from (3.6.2) and (3.6.3),

$$u_1 = c_1 t^{a_2},\ u_2 = c_2 t^{a_1},\ u_3 = c_3 t^{a_o} \quad \text{in the equilateral case,} \tag{3.6.4}$$

$$u_1 = c_1 t^{b_o},\ u_2 = c_2 t^{b_1} \quad \text{in the collinear case.} \tag{3.6.5}$$

We have to examine the possibility of multiple eigen-values. Since in the collinear case, the possible double roots b_2, b_3 have positive real parts and the solution depends only on the eigen-values with negative real parts, it follows that the only solutions are those given by (3.6.5) whenever $\lambda_k \neq \sum\limits_{r=1}^{p} g_r \lambda_r$. In the equilateral case, the only possible double eigen-value is $a_1 = a_2$. Then $\lambda_1 = \sum\limits_{r=1}^{3} g_r \lambda_r$ is satisfied with $g_1 = 0$, $g_2 = g_3 = 0$.

We now consider the case in which for some k, we have $\lambda_k = \sum\limits_{r=1}^{p} g_r \lambda_r$, $g_1 + \ldots + g_p \geq 2$. We take the equilateral case first. We have to discuss the possibility of the existence of non-negative integers g_1, g_2, g_3 with $g_1 + g_2 + g_3 \geq 2$ such that $-\lambda_k = g_1 a_2 + g_2 a_1 + g_3 a_o$, where $-\lambda_1 = a_2$, $-\lambda_2 = a_1$, $-\lambda_3 = a_o (a_o > a_1 \geq a_2)$. Since the polynomial $V_1 \equiv 0$ always, it is enough to consider only the two cases $k = 2, 3$ corresponding to $-\lambda_2 = a_1$ and $-\lambda_3 = a_o$.

If $k = 2$, then $-\lambda_2 = a_1$ and $g_2 = g_3 = 0$, so that $V_2(u_1) = \alpha u_1^{g_1}$. We 234
have $a_1 = g_1 a_2$, or $\frac{a_1}{a_2} = g_1 \geq 2$. Denoting this integer by h, we have

$$h = \frac{a_1}{a_2} = \frac{\sqrt{13 + w} - 1}{\sqrt{13 - w} - 1} \text{ where } w = 12\sqrt{1 - 3a}. \tag{3.6.6}$$

For each integer $h = 2, 3, \ldots$, we can determine w and a from this equation and $V_2(u_1) = \alpha u_1^h$.

If $k = 3$, then $-\lambda_3 = a_o = g_1 a_2 + g_2 a_1$ with $g_1 + g_2 \geq 2$, since in this case $g_3 = 0$. We observe that $2a_1 > a_o$, so $g_2 = 0$ or 1. Hence, either $g_1 \geq 2$ (if $g_2 = 0$) or $g_1 \geq 1$ (if $g_2 = 1$). We show that g_2 cannot be 1, so $g_1 \geq 2$ necessarily. In fact, since we see that $a_2 + a_1 > a_o$, we have $g_2 = 0$ and so $g_1 \geq 2$. Hence $a_o = g_1 a_2$ or $\dfrac{a_o}{a_2} = g_1 \geq 2$. Denoting this integer by g we have

$$g = \frac{a_o}{a_2} = \frac{4}{\sqrt{13 - w} - 1} \text{ where } w = 12\sqrt{1 - 3a}. \tag{3.6.7}$$

Again we can determine w and a from (3.6.7) for each $g \geq 2$. In this case $V_3(u_1, u_2) = \beta u_1^g$.

We next prove that only one of the possibilitie

$$V_2(u_1) = \alpha u_1^h,\ V_3(u_1, u_2) = 0;\ V_2(u_1) = 0,\ V_3(u_1, u_2) = \beta u_1^g,$$

can occur. (It cannot happen that $V_2(u_1) = \alpha_1^h$ and $V_3(u_1, u_2) = \beta u_1^g$). In order to prove this, we show that (3.6.6) and (3.6.7) cannot be satisfied simultaneously for integers $g, h \geq 2$. We have, from (3.6.6) and (3.6.7),

$$\sqrt{13 - w} = 1 + \frac{4}{g},\ \ \sqrt{13 + w} = 1 + \frac{4h}{g}.$$

235 Eliminating w by squaring and adding we get

$$26g^2 = (g + 4)^2 + (g + 4h)^2.$$

So it is enough to prove that this diophantine equation does not have integer solutions $g, h \geq 2$. We can write this as $(25g - 4)^2 = 416 + 25(g + 4h)^2$ from which we have

$$\begin{aligned} 416 = (25g - 4)^2 - 25(g + 4h)^2 &= (30g + 20h - 4)\,(20g - 20h - 4) \\ \text{or} \qquad (15g + 10h - 2)\,(5g - 5h - 1) &= 52. \end{aligned}$$

Setting $p = 15g + 10h - 2$, $q = 5g - 5h - 1$, we have to show that there are no integers p, q such that

$$pq = 52,\ \ p - 3q - 1 = 25h, \tag{3.6.8}$$

By definition of p, $p > 0$ and since $pq = 52$, $q > 0$. The only integer factorisations pq of 52, with $q < p$, are $p = 52$, $q = 1$, $p = 26$, $q = 2$ and $p = 13$, $q = 4$. It is easy to check that none of these factorisations satisfies (3.6.8) with integer $h \geq 1$. This proves that only one of the exceptional cases

$$V_2(u_1) = \alpha u_1^h, \ V_3(u_1, u_2) = 0; \tag{3.6.9}$$

$$V_2(u_1) = 0, \ V_3(u_1, u_2) = \beta u_1^g, \tag{3.6.10}$$

can occur.

Suppose that the possibility (3.6.9) holds. Then $u_1 = c_1 e^{\lambda_1 s} = c_1 t^{a_2}$, **236**
since $\lambda_1 = -a_2$ and $e^{-s} = t$, $c_1 = u_1(0)$. Inserting this in $u_2' = \lambda_2 u_2 + V_2(u_1)$ we get

$$u_2' = -a_1 u_2 + \alpha u_1^{g_1} = -a_1 u_2 + \alpha c_1^h e^{\lambda_1 hs} = -a_1 u_2 + \alpha c_1^h e^{-a_1 s},$$

so $(u_2 e^{a_1 s})' = \alpha c_1^h$, and integrating from 0 to s and putting $u_2(0) = c_2$,

$$u_2 e^{a_1 s} = c_2 + \alpha c_1^h s, \text{ or } u_2 = (c_2 + \alpha c_1^h s) e^{-a_1 s}.$$

Thus, if (3.6.9) holds, then

$$u_1 = c_1 t^{a_2}, u_2 = (c_2 - \alpha c_1^h \log t) t^{a_1}, u_3 = c_3 t^{a_o}. (h \text{ integer } \geq 2)$$

Similarly, if (3.6.10) holds, we have

$$u_1 = c_1 t^{a_2}, u_2 = c_2 t^{a_1}, u_3 = (c_3 - \beta c_1^g \log t) t^{a_o}. (g \text{ integer } \geq 2)$$

In the equilateral case, if we have a double root $a_1 = a_2$ (when $m_1 = m_2 = m_3$), we have $h = 1$. So $V_2(u_1) = \alpha u$ and hence

$$u_1 = c_1 t^{a_2}, u_2 = (c_2 - \alpha c_1 \log t) t^{a_1}, u_3 = c_3 t^{a_o}.$$

The value of α will be just e_2(= 0 or 1) since $\lambda_1 = -a_2 = -a_1 = \lambda_2$.

Finally we consider the collinear case. In this case there is only condition $\lambda_k = \sum_{r=1}^{2} g_r \lambda_r$, i.e. $b_1 = j b_o$, j integer ≥ 2, because $\lambda_1 = -b_o$,

$\lambda_2 = -b_1$ and $0 < -\lambda_1 < -\lambda_2$ this gives $b = j^2 + \dfrac{j-3}{2}$. Moreover, we have $V_2(u_1) = \gamma u_1^j$ so that we have

$$u_1 = c_1 2^{\lambda_1 s} = c_1 e^{-b_o s} = c_1 t^{b_o}, u_2 = (c_2 + \gamma c_1^j s)e^{\lambda_2 s} = (c_2 - \gamma c_1^j \log t)t^{b_1}.$$

We collect together the results in the two cases. **237**

Equilateral case.

(i) If $\dfrac{a_o}{a_2}$ and $\dfrac{a_1}{a_2}$ are not integers $u_1 = c_1 t^{a_2}$, $u_2 = c_2 t^{a_1}$, $u_3 = c_3 t^{a_o}$, $u_k = 0$, $k = 4, 5, 6$.

(ii) If $\dfrac{a_o}{a_2}$ = an integer $g \geq 2$,

$$u_1 = c_1 t^{a_2}, u_2 = c_2 t^{a_1}, u_3 = (c_3 - \alpha c_1^g \log t)t^{a_o}, u_k = 0, k = 4, 5, 6,$$

(iii) $u_1 = c_1 t^{a_2}$, $u_2 = (c_2 - \beta c_1^h \log t)t^{a_1}$, $u_3 = c_3 t^{a_o}$, $u_k = 0$, $k = 4, 5, 6$, where $\dfrac{a_1}{a_2}$ = an integer $h \geq 1$.

Collinear case.

(i) If $\dfrac{b_o}{b_1}$ is not integral then

$$u_1 = c_1 t^{b_o},\ u_2 = c_2 t^{b_1},\ u_k = 0,\ k = 3, 4, 5, 6.$$

(ii) If $\dfrac{b_1}{b_o}$ = an integer $j \geq 2$,

$$u_1 = c_1 t^{b_o}, u_2 = (c_2 - \gamma c_1^j \log t)t^{b_1}, u_k = 0, k = 3, 4, 5, 6.$$

The constants of integration c_k have small absolute values.

By taking the inverse of the transformation $u_k = y_k - F_k(y_1, \ldots, y_n)$, $k = 1, \ldots, m$, we have $y_k = u_k + G_k(u_1, \ldots, u_n)$, $k = 1, \ldots, m$, where G_k are power series, with coefficients complex in general, starting with quadratic terms and converging for small $|u_k|$. In our case, $m = 6$ and

238 $n = 3$ or 2 according us we are in the equilateral or in the collinear case. So we see after inserting the above solutions u_k, that $y_1, \ldots, y_6$ are power-series, without constant terms, in u_1, u_2, u_3 in the equilateral case, and in u_1, u_2 in the collinear case. Since y_k are obtained from δ_k by a linear transformation $\delta = Cy$, $|C| \neq 0$, it follows that $\delta_1, \ldots, \delta_6$ are power-series without constant terms, in u_1, u_2, in the collinear case and in u_1, u_2, u_3 in the equilateral case; these power-series converge for small $|u_k|$, and hence for sufficiently small t. Since the power-series do not have constant terms and $u_k(t) \to 0$ as $t \to 0$, it follows that $\delta_k \to 0$ as $t \to 0$.

We recall that $p_k^* = \overline{p}_k + \delta_k$, $q_k^* = \overline{q}_k + \delta_{k+3}$, $k = 1, 2, 3$, where $\overline{p}_k, \overline{q}_k$ are uniquely determined by the masses. We also have $\overline{p}_1 > 0$. Since $q_4 = 0$ for a collision orbit, we have $\delta_7 = 0$. We have by definition $p_4 = p_4^* = \delta_8$ and δ_8 satisfies the differential equation

$$\delta_8' = \sum_{l=1}^{6} a_{8l}\delta_l + \varphi_8(\delta_1, \ldots, \delta_6), \tag{3.6.11}$$

where a_{8l} are real constants determined uniquely by the masses and φ_8 is a power-series with real coefficients, again uniquely determined by the masses, and starting with quadratic terms. Hence the right side of (3.6.11) is a power-series $Q = Q(\delta_1, \ldots, \delta_6)$ with real coefficients and without constant terms. We have determined $\delta_1, \ldots, \delta_6$ as power-series in $u_k (k = 3$ or $2)$ without constant terms and converging for small $|u_k|$. Hence Q is also a power-series in u_k without constant term and converging for small $|u_k|$. So $p_4 = \delta_8$ can be obtained by integrating Q from 0 239
to s. In order to study the behaviour of p_4 as $t \to 0$ (or, equivalently, as $s \to \infty$), we have to prove the convergence, as $s \to \infty$, of the integral

$$\int_{o}^{s} Q(\delta_1, \ldots, \delta_6) ds.$$

(We do this only in the case in which the relation $\lambda_k \neq \sum_{r=1}^{p} g_r \lambda_r$ holds; the discussion in the other case is similar). A typical term in Q is

$\beta_{h_1,\ldots,h_n} u_1^{h_1} \ldots u_n^{h_n}$, where $n = 3$ or 2 according as we are in the equilateral or collinear case. In the equilateral case, u_1, u_2, u_3 are given by (3.6.4) when a_o, a_1, a_2 are distinct and so a typical term of Q is $\beta_{h_1 h_2 h_3} c_1^{h_1} c_2^{h_2} c_3^{h_3} e^{-(h_1 a_2 + h_2 a_1 + h_3 a_o)s}$. This term is integrable in $0 \leq s \leq \infty$ and we have

$$\int\limits_0^s e^{-(h_1 a_2 + h_2 a_1 + h_3 a_o)s} ds = \frac{e^{-(h_1 a_2 + h_2 a_1 + h_3 a_o)s} - 1}{-(h_1 a_2 + h_2 a_1 + h_3 a_o)}$$

and this tends to $(h_1 a_2 + h_2 a_1 + h_3 a_o)^{-1}$ as $s \to \infty$. Since the constants c_1, c_2, c_3 have small absolute values, so have u_1, u_2, u_3. Since Q converges uniformly for small $|u_k|, k = 1, 2, 3$, we can integrate the power-series and obtain

$$p_4 = \delta_8 = \int^s Q(\delta_1, \ldots, \delta_6) ds + \overline{p}_4$$

where $\overline{p}_4$ is a constant of integration, and so we have $p_4 = \overline{p}_4 +$ a unique
240 power-seires in u_1, u_2, u_3 withtout constant term. The power-series converges for small $|u_k|$ and so for small t. This proves that p_4 tends to a finite limit $\overline{p}_4$ as $t \to 0$.

Since the differential equations of motion remain invariant under an orthogonal transformation of coordinates, we can perform a fixed orthogonal transformation in the plane of motion, so that $\overline{p}_4$ becomes 0 in the new coordinate system. In other words, we may assume that $p_4 \to 0$ as $t \to 0$. A similar argument can be carried out in the case of double eigen-values, and also in the collinear case, and we see that p_4 tends to a finite limit, which may be assumed to be 0, as $t \to 0$, in all cases.

We shall now go back to the original coordinate system and find x_k, y_k, $k = 1, 2, 3$, the coordinates of P_1, P_2, P_3. First we have

$$\begin{aligned}
p_k &= p_k^* t^{2/3} = (\overline{p}_k + \delta_k) t^{2/3}, \overline{p}_k > 0, \\
q_k &= q_k^* t^{-1/3} = (\overline{q}_k + \delta_{k+3}) t^{-1/3}, k = 1, 2, 3, \\
p_4 &= p_4^* = \delta_8, q_4 = q_4^* = 0.
\end{aligned}$$

Hence p_k^* are power-series in u_k, $k = 1, 2, 3$ or $k = 1, 2$ according as we are in the equilateral or collinear case, which converge for small $|u_k|$. And $p_k = p_k^*(u_l)t^{2/3}$, $q_k = q_k^*(u_l)t^{-1/3}$, $k = 1, 2, 3$. We can also calculate $c = \cos p_4$ and $s = \sin p_4$ as power-series in u_l which converge for small $|u_l|$. Since

$$\xi_1 = p_1 c, \xi_2 = p_1 s, \xi_3 = p_2 c - p_3 s, \xi_4 = p_2 s + p_3 c,$$
$$\eta_1 = q_1 c - q_o s, \eta_2 = q_1 s + q_o c,\ \eta_3 = q_2 c - q_3 s, \eta_4 = q_2 s + q_3 c,$$

we see that the relative coordinates (ξ_1, ξ_2) and (ξ_3, ξ_4) of P_1 and P_2 with **241**
respect to P_3 and the corresponding components of momenta $\eta_1, \ldots, \eta_4$, are also given by convergent power-series in u_k:

$$\xi_k = \xi_k^*(u_1, u_2, u_3)t^{2/3},\ \eta_k = \eta_k^*(u_1, u_2, u_3)t^{-1/3}$$

in the equilateral case, and

$$\xi_k = \xi_k^*(u_1, u_2)t^{2/3},\ \eta_k = \eta_k^*(u_1, u_2)t^{-1/3}$$

in the collinear case. We can now go back to the absolute coordinates x_k, y_k. If v denotes any of the six coordinates $x_k, y_k, k = 1, 2, 3$, then we have

$$v = t^{2/3}P(u_1, u_2, u_3) \text{ and } v = t^{2/3}P(u_1, u_2)$$

in the equilateral and collinear cases respectively, where P is a power-series convergent for small $|u_k|$, and hence for small $|c_k|$. The components of momenta also have power-series expansions

$$w = t^{-1/3}H(u_1, u_2, u_3), w = t^{-1/3}H(u_1, u_2),$$

H converging for small $|u_k|$ and hence for small $|c_k|$.

We consider now the manifold of all the collision orbits. Corresponding to different values of the real parameters c_1, c_2, c_3 in the equilateral case and c_1, c_2 in the collinear case, we obtain different collision orbits in a neighbourhood of $t = 0$. Let us determine the dimensions of this manifold in the neighbourhood of $t = 0$. The coordinates **242**
$x_k, y_k, k = 1, 2, 3$, being power-series in u_k ($k = 1, 2, 3$ or $k = 1, 2$), we

have three real parameters in the equilateral case and two real parameters in the collinear case. We have proved that the motion takes place in a fixed plane, chosen as the (x, y)-plane in our discussion. Its position is determined by the three angles of a normal to it: this involves three extra independent real parameters in both the limiting cases. We have assumed that the centre of gravity remains fixed at the origin, and the centre of gravity integrals involve six real parameters. Finally, the points on a orbit are parametrised by the real variable t. Thus in the equilateral case we have 13 independent real parameters, and in the collinear case 12. Since the coordinate functions are regular analytic functions of these parameters, we conclude that the manifold of all collision orbits is a real analytic manifold, of dimension 13 in the equilateral case and 12 in the collinear case, in a neighbourhood of $t = 0$. Further, in the collinear case there are three distinct orderings of the points P_1^*, P_2^*, P_3^* and hence, corresponding to these, there are three distinct real 12-dimensional analytic manifolds. (In the case of a simple collision we have seen that we have power-series expansions for the coordinates in the variable $t^{1/3}$; the manifold of collision orbits is there a real analytic manifold of dimension 16 in the neighbourhood of a simple collision).

We remark that since our solutions are described only near $t = 0$, the
243 above description of the collision orbits is purely local. It is not possible to describe the manifold of collision orbits in the large, that is, for all t, by our method.

We consider the nature of the singularity when there is a general collision at $t = 0$. Since the coordinates are power-series in u_1, u_2, u_3 or in u_1, u_2, the nature of the singularity depends on the arithmetical nature of the eigen-values a_o, a_1, a_2 in the equilateral case and b_o, b_1 in the collinear case. If a_2, a_1 in the equilateral case and b_1 in the collinear case are rational, then we have an algebraic branch point at $t = 0$, so that the solutions of the three-body problem can be uniformised in a neighbourhood of $t = 0$. If a_2, a_1 are irrational and $c_1, c_2 \neq 0$ in the equilateral case, and b_1 irrational and $c_2 \neq 0$ in the collinear case, we have an essential singularity at $t = 0$. In this case it is not possible to continue the solutions analytically beyond the general collision.

□