



---

Soviet-era science, translated into English

# V. I. ARNOLD

1965

SovietRxiv

---

View the original and related papers at <https://sovietrxiv.org/items/ru-196501.13533>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

**Abstract**

**Full Text**

**V. I. ARNOLD**

**CONDITIONS FOR APPLICABILITY AND AN ERROR ESTIMATE FOR THE AVERAGING METHOD FOR SYSTEMS WHICH, IN THE COURSE OF EVOLUTION, PASS THROUGH RESONANCES**

*(Presented by Academician A. N. Kolmogorov, 7 X 1964)*

**§ 1. Behavior of solutions of systems of the form**

$$\dot{\varphi} = \omega(I; \varepsilon) + \varepsilon f(I, \varphi; \varepsilon),$$

$$\dot{I} = \varepsilon F(I, \varphi; \varepsilon), \tag{1}$$

$$\varphi = \varphi_1, \dots, \varphi_k; \quad I = I_1, \dots, I_l$$

(where  $\varphi \pmod{2\pi}$  are angles;  $\varepsilon \ll 1$ ; a dot denotes differentiation with respect to time  $t$ ; the functions  $\omega, f, F$  are analytic for  $I \in G$ ,  $|\operatorname{Im} \varphi| < \rho$ ,  $|\varepsilon_0| < \varepsilon_0$  (their dependence on  $\varepsilon$  will not be indicated below);  $G$  is a complex compact domain) is usually studied by the “averaging method,” i.e., by replacing (1) with the averaged system

$$\dot{J} = \varepsilon \bar{F}(J), \quad \bar{F}(J) = (2\pi)^{-k} \iint F(J, \varphi; 0) d\varphi. \tag{2}$$

Although the terms discarded in averaging,  $\varepsilon F = \varepsilon F - \varepsilon \bar{F}$ , are of the same order of magnitude as those retained, it is believed that over a time  $t \sim 1/\varepsilon$  the difference between the exact and averaged solutions with identical initial conditions,  $|I(t) - J(t)|$ , remains small. Indeed, in the one-frequency case ( $k = 1$ ) one easily obtains<sup>1</sup> the estimate

$$|I(t) - J(t)| < C_1 \varepsilon \quad \text{for } 0 < t < 1/\varepsilon.$$

Here and below  $C_1, \dots, C_{22}$  are sufficiently large constants independent of  $\varepsilon, K, N, x$ .

In the present note a two-frequency system ( $k = 2$ ) is considered. We shall indicate a condition sufficient for the smallness of  $|I - J|$ , and obtain the estimates

$$C_2^{-1}\sqrt{\varepsilon} < |I(t) - J(t)| < C_3\sqrt{\varepsilon} \ln^2(1/\varepsilon).$$

**§ 2. First of all we give an example showing that without additional assumptions averaging may lead to incorrect results.**

**Example 1.** Consider the system

$$\dot{\varphi}_1 = I_1, \quad \dot{\varphi}_2 = I_2, \quad \dot{I}_1 = \varepsilon, \quad \dot{I}_2 = \varepsilon a \cos(\varphi_1 - \varphi_2) \quad (a > 1).$$

The averaged equations are  $\dot{J}_1 = \varepsilon$ ,  $\dot{J}_2 = 0$ . Let  $I_1(0) = I_2(0) = J_1(0) = J_2(0) = 1$ ,  $\varphi_1(0) = 0$ ,  $\varphi_2(0) = \arccos(1/a)$ . The exact solution  $I_1(t) = I_2(t) = 1 + \varepsilon t$  after the time  $t = 1/\varepsilon$  has lost any connection with the averaged solution  $J_1(t) = 1 + \varepsilon t$ ,  $J_2(t) = 1$ .

Returning to system (1),  $k = 2$ , suppose that  $\omega_2(I) \neq 0$ . Introduce the frequency ratio  $\lambda(I) = \omega_1/\omega_2$ .

**Condition A.** Suppose that  $C_4^{-1}\varepsilon < |\dot{\lambda}| < C_4\varepsilon$ , i.e., that the quantity

$$A(I, \varphi) = \left( \frac{\partial \omega_1}{\partial I} F \right) \omega_2 - \left( \frac{\partial \omega_2}{\partial I} F \right) \omega_1$$

does not vanish for any  $\varphi$ , if  $I \in G$ .

Under condition A the system cannot get stuck at any resonance. In example 1 condition A is violated:  $A = I_2 - I_1 \cos(\varphi_1 - \varphi_2)$  changes sign for  $I_1 = I_2$ , if  $a > 1^*$ .

**§ 3. Theorem 1.** *If condition A is fulfilled, then the estimate*

$$|I(t) - J(t)| < C_3\sqrt{\varepsilon} \ln^2(1/\varepsilon) \quad \text{for all } 0 \leq t \leq 1/\varepsilon \quad (3)$$

*is valid.*

Roughly speaking, theorem 1 estimates the difference between the solutions of the exact and averaged systems by a quantity  $\sqrt{\varepsilon}$ . The following example shows that, in general,  $|I(t) - J(t)| > C_2^{-1}\sqrt{\varepsilon}$ .

**Example 2.** Consider the system

$$\dot{\varphi}_1 = I_1 + I_2, \quad \dot{\varphi}_2 = I_2; \quad \dot{I}_1 = \varepsilon, \quad \dot{I}_2 = \varepsilon \cos(\varphi_1 - \varphi_2).$$

Let  $\varphi_1(0) = \varphi_2(0) = I_1(0) = I_2(0) - 1 = 0$ . Condition A is fulfilled for  $I_1 < I_2$ . In the averaged system  $J_2(t) \equiv 1$ . In the exact solution

$$I_2(T) - 1 = \varepsilon \int_0^T \cos \varepsilon \frac{t^2}{2} dt = \sqrt{2\varepsilon} \int_0^\tau \cos x^2 dx, \quad \tau = \sqrt{\varepsilon/2} T.$$

For  $T = 1/\varepsilon$ , evidently,

$$I_2(T) - J_2(T) = I_2(T) - 1 > C_2^{-1} \sqrt{\varepsilon} **.$$

Analyzing example 2, it is easy to notice that the resonance  $\omega_1 = \omega_2$  has a scattering effect on a pencil of trajectories that initially differ only in the phases  $\varphi$ . The spread of the quantity  $I_2$  after passage through the resonance is of order  $\sqrt{\varepsilon}$ .

The idea of the proof of theorem 1 consists in dividing the space  $I$  into two parts: a finite (of order  $\ln^2(1/\varepsilon)$ ) number of resonant zones of width  $K$  and a nonresonant part. During the time spent in the resonant zones a discrepancy  $|I - J|_p \sim K$  accumulates (neglecting logarithms). In the nonresonant region we construct new variables  $P$  satisfying the conditions

$$|P - I| \sim \varepsilon/K, \quad |\dot{P} - \varepsilon \bar{F}(P)| \sim \varepsilon^2/K^2.$$

From these estimates we derive that

$$|P - J| \sim \varepsilon/K.$$

Thus,

$$|I - J|_{\text{nr}} \leq |I - P| + |P - J| \sim \varepsilon/K.$$

Consequently,

$$|I - J| \leq |I - J|_p + |I - J|_{\text{nr}} \sim K + \varepsilon/K.$$

For  $K \sim \sqrt{\varepsilon}$  we obtain (3).

§ 4. **Estimates.** Let  $N > 1 > K > 0$ . Denote by  $G_N$  the set of points  $I$  of the domain  $G$  at which

$$(\omega, n) = \omega_1 n_1 + \omega_2 n_2 \neq 0$$

for integer  $n_1, n_2$ ;

$$0 < |n| = |n_1| + |n_2| < N.$$

Denote by  $G_{K,N}$  the set of points belonging to  $G_N$  together with a neighborhood of radius  $K$ .

From condition A it follows that  $d(\omega, n) \neq 0$  when  $(\omega, n) = 0$ . Therefore in  $G_{K,N}$

$$|(\omega, n)| > C_5^{-1} K, \quad 0 < |n| < N. \quad (4)$$

We denote the complement of  $G_{K,N}$  by  $R_{K,N} = G \setminus G_{K,N}$ . Let  $I(t), \varphi(t)$  ( $0 \leq t \leq 1/\varepsilon$ ) be a solution of system (1), with  $I(t) \in G$ . The interval  $0 \leq t \leq 1/\varepsilon$  is divided into two parts:  $g_{K,N}$  (where  $I(t) \in G_{K,N}$ ) and  $r_{K,N}$  (where  $I(t) \in R_{K,N}$ ). Let  $K < C_6^{-1}$ . From condition A it follows

---

\* Example 1 shows that condition A cannot be replaced by the analogous condition  $\bar{A}$  on the averaged system (2), as is proposed in (2). However, it is possible that under condition A inequality (3) is valid for the majority of initial conditions. Namely, this is the case in example 1, which is readily integrated to the end:  $\ddot{q} = -\partial U/\partial q$ ,  $U = -\varepsilon(q - a \sin q)$ , where  $q = \varphi_1 - \varphi_2$ . See also (3, 4).

\*\* Example 2 refutes the assertion, made in (2), that  $|I - J| < C\varepsilon$ . Probably, in the general case

$$|I - J| > C_2^{-1} \sqrt{\varepsilon} \ln^2(1/\varepsilon).$$

**Lemma 1.** The set  $r_{K,N}$  consists of no more than  $C_7 N^2$  intervals. The length of each of them does not exceed  $C_8 K/\varepsilon$ .

**Lemma 2.** Let  $\alpha < t < \beta$  be one of the intervals constituting  $g_{K,N}$ . If  $\alpha + x \leq t \leq \beta - x$ , then  $I(t) \in G_{K(x),N}$ , where  $K(x) = K + C_9^{-1} \varepsilon x$ .

Denote by  $J(J_0, t_0; t)$  the solution of system (2) with  $J(t_0) = J_0$ . It is obvious that

**Lemma 3.** For  $|t - t_0| < 1/\varepsilon$  we have

$$|J(J_0, t_0; t) - J(J'_0, t_0; t)| < C_{10} |J_0 - J'_0|.$$

Let  $|\dot{x}| \leq a|x| + b(t)$ ,  $|x(0)| < c$ ;  $a, b, c \geq 0$ . It is easily proved that

**Lemma 4.**

$$|x(t)| \leq \left[ c + \int_0^t b(t) dt \right] e^{at}.$$

By the usual methods (see § 6) the main result is proved:

**Lemma 5.** There exist functions  $P = I + S(I, \varphi)$ ,  $S(I, \varphi + 2\pi) = S(I, \varphi)$ , and constants  $C_{11} - C_{14}$ , independent of  $\varepsilon, K$ , such that

$$|\dot{P} - \varepsilon \bar{F}(P)| < C_{11} \varepsilon^2 / K^2, \quad |P - I| < C_{12} \varepsilon / K \quad (5)$$

for  $|\operatorname{Im} \varphi| \leq 0.5\rho$ ;  $I \in G_{K,N}$ ;  $N = C_{13} \ln(1/\varepsilon)$ ;  $|\varepsilon| \leq C_{14}^{-1} K^2$ .

§ 5. **Proof of Theorem 1.** Let  $0 < \varepsilon < C_{14}^{-1} C_6^{-2}$ ,  $K = \sqrt{C_{14} \varepsilon} < C_6^{-1}$ . Applying Lemma 5, form  $P(I, \varphi)$  and  $G_{K,N}$ . Denote the consecutive intervals constituting  $g_{K,N}$  by  $[t_r^I, t_r^II]$  ( $r = 1, 2, \dots$ ; for definiteness,  $t_1^I = 0 \in g_{K,N}$ ;  $1/\varepsilon \in r_{K,N}$ ).

Introduce the notation  $\alpha = \text{I, II}$ ;  $I(t_r^\alpha) = I_r^\alpha$ ;  $J(I_r^\alpha, t_r^\alpha; t) = J_r(t)$ ;  $P(I(t), \varphi(t)) = P(t)$ . From Lemma 3, taking into account that  $J_r(t_r^{\text{I}}) = I_r^{\text{I}}$ , we obtain

$$\begin{aligned} |I(t) - J(t)| &\leq \sum_r |J_{r+1}(t) - J_r(t)| \leq C_{10} \sum_r |J_{r+1}(t_{r+1}^{\text{I}}) - J_r(t_{r+1}^{\text{I}})| \\ &\leq C_{10} \sum_r \{|I_{r+1}^{\text{I}} - I_r^{\text{II}}| + |I_r^{\text{II}} - J_r(t_r^{\text{II}})| + |J_r(t_r^{\text{II}}) - J_r(t_{r+1}^{\text{I}})|\}. \end{aligned} \quad (6)$$

From Lemma 1 it follows that

$$|I_{r+1}^{\text{I}} - I_r^{\text{II}}| + |J_r(t_{r+1}^{\text{I}}) - J_r(t_r^{\text{II}})| < C_{15}K. \quad (7)$$

From Lemma 5 we find, for  $t_r^{\text{I}} \leq t \leq t_r^{\text{II}}$ ,

$$\begin{aligned} a &= C_{16}\varepsilon, \quad b = |\dot{P} - \varepsilon\bar{F}(P)|, \\ |\dot{P} - \dot{J}_r| &\leq a|P - J_r| + b. \end{aligned} \quad (8)$$

We estimate the quantity  $b$  with the aid of Lemmas 2 and 5:

$$b(t) < C_{11}\varepsilon^2/(K + C_9^{-1}\varepsilon x)^2 \quad \text{for } t_r^{\text{I}} + x \leq t \leq t_r^{\text{II}} - x.$$

Consequently,

$$\int b(t) dt < C_{17}\varepsilon/K \quad \text{for } t_r^{\text{I}} \leq t \leq t_r^{\text{II}}.$$

According to (5),

$$|P(t_r^{\text{I}}) - J(t_r^{\text{I}})| = |P(t_r^{\text{I}}) - I_r^{\text{I}}| < c = C_{12}\varepsilon/K.$$

Applying now Lemma 4 to inequality (8), we find

$$|P(t_r^{\text{II}}) - J_r(t_r^{\text{II}})| \leq e^{C_{16}}(C_{12} + C_{17})\varepsilon/K < C_{18}\varepsilon/K. \quad (9)$$

From (5) and (9) it follows that

$$|I_r^{\text{II}} - J_r(t_r^{\text{II}})| \leq |I_r^{\text{II}} - P(t_r^{\text{II}})| + |P(t_r^{\text{II}}) - J_r(t_r^{\text{II}})| < (C_{12} + C_{18})\varepsilon/K. \quad (10)$$

Combining (6), (7), (10), and Lemma 1, we find, for  $0 \leq t \leq 1/\varepsilon$ ,

$$|I(t) - J(t)| < C_{10}C_7(C_{13} \ln(1/\varepsilon))^2 [C_{15}K + (C_{12} + C_{18})\varepsilon/K].$$

For  $K = \sqrt{C_{14}\varepsilon}$  the right-hand side is less than  $C_3\sqrt{\varepsilon} \ln^2(1/\varepsilon)$ , as was required.

§ 6. Proof of Lemma 5. Let

$$S = \sum S_n e^{i(n,\varphi)}; \quad S_n = \frac{i\varepsilon F_n}{(\omega, n)}; \quad 0 < |n| < N, \quad (11)$$

where  $F(I, \varphi) = \bar{F}(I) + \tilde{F}(I, \varphi)$  is expanded in a Fourier series:  $\tilde{F} = \sum F_n e^{i(n,\varphi)}$  ( $|n| > 0$ ), and  $[F]_N = \sum F_n e^{i(n,\varphi)}$  ( $0 < |n| < N$ ). Then  $\dot{P} = \varepsilon \bar{F}(P) + \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5$ , where  $\Sigma_1 = [\varepsilon \tilde{F}]_N + \frac{\partial S}{\partial \varphi} \omega \equiv 0$  (see (11));  $\Sigma_2 = \varepsilon \bar{F}(I) - \varepsilon \bar{F}(P)$ ;  $\Sigma_3 = \varepsilon \tilde{F} - [\varepsilon \tilde{F}]_N$ ;  $\Sigma_4 = \frac{\partial S}{\partial \varphi} \varepsilon F$ ;  $\Sigma_5 = \frac{\partial S}{\partial \varphi} \varepsilon f$ .

For  $|\operatorname{Im} \varphi| < 0.9\rho$ ,  $I \in G_{0.1K;N}$ , in view of (4), we have (cf. (5))

$$|S| < C_{12}\varepsilon/K, \quad |\partial S/\partial \varphi| < C_{12}\varepsilon/K, \quad |\partial S/\partial I| < C_{12}\varepsilon/K^2. \quad (12)$$

If  $I \in G_{K,N}$ ,  $|\varepsilon| < C_{14}^{-1}K^2$ , then the entire segment  $IP \subset G_{0.1K;N}$ . Therefore, for  $I \in G_{K,N}$  and  $|\operatorname{Im} \varphi| < 0.5\rho$ , we have, according to (12),

$$|\Sigma_2| < \varepsilon \left| \frac{\partial \bar{F}}{\partial I} S \right| < c_{19}\varepsilon^2/K; \quad |\Sigma_4| < C_{20}\varepsilon^2/K^2, \quad |\Sigma_5| < C_{21}\varepsilon^2/K.$$

In view of the analyticity of  $\tilde{F}$  for  $|\operatorname{Im} \varphi| < \rho$ , from  $N = C_{13} \ln(1/\varepsilon)$ , with sufficiently large  $C_{13} = C_{13}(\rho)$ , it follows that  $|\Sigma_3| < C_{22}\varepsilon^2$ . Thus,  $|\dot{P} - \varepsilon \bar{F}(P)| < C_{11}\varepsilon^2/K^2$  (where  $C_{11} = C_{19} + C_{20} + C_{21} + C_{22}$ ), as was required to prove.

The present note was prompted by the erroneous work (2). The author expresses gratitude to A. M. Molchanov, who pointed out this work.

Moscow State University  
named after M. V. Lomonosov

Received  
1 X 1964

## References

1. N. N. Bogolyubov, Yu. A. Mitropolskii, *Asymptotic Methods in the Theory of Nonlinear Oscillations*, Moscow, 1958.
2. A. M. Molchanov, in: *Problems of the Motion of Artificial Celestial Bodies*, Moscow, 1963, p. 42.
3. D. V. Anosov, *Izv. AN SSSR, Ser. Matem.*, 24, 721 (1960).
4. T. Kasuga, *Proc. Japan Acad.*, 37, No. 7, 366, 372, 377 (1961).

5. V. I. Arnold, *UMN*, 18, No. 5, 13 (1963); 18, No. 6, 91 (1963).

*Note: Figure translations are in progress. See original paper for figures.*

*Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.*